315930

Studia Scientiarum Mathematicarum Hungarica

31/1996

13

14.

EDITOR-IN-CHIEF

D. SZÁSZ

VOLUME 31
NUMBERS 1–3
1996

AKADÉMIAI KIADÓ, BUDAPEST

STUDIA SCIENTIARUM MATHEMATICARUM HUNGARICA

A QUARTERLY OF THE HUNGARIAN ACADEMY OF SCIENCES

Studia Scientiarum Mathematicarum Hungarica publishes original papers on mathematics mainly in English, but also in German, French and Russian. It is published in yearly volumes of four issues (mostly double numbers published semiannually) by

AKADÉMIAI KIADÓ
Publishing House of the Hungarian Academy of Sciences
H-1117 Budapest, Prielle Kornélia u. 19-35

Manuscripts and editorial correspondence should be addressed to

J. Merza
Managing Editor

P.O. Box 127
H-1364 Budapest

Tel.: +36 1 118 2875 Fax: +36 1 117 7166
e-mail: merza @ math-inst.hu

Subscription information

Orders should be addressed to

AKADÉMIAI KIADÓ
P.O.Box 245
H-1519 Budapest

For 1996 volumes 31-32 are scheduled for publication. The subscription price is $ 104.00 each, air delivery plus $ 20.00 per volume.

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 1–13

LIMIT LAWS FOR THE MAXIMAL AND MINIMAL INCREMENTS OF THE POISSON PROCESS

P. AUER and K. HORNIK

Dedicated to Professor P. Révész for his 60th birthday

Abstract

We give a partial survey of strong laws for maximal and minimal increments of homogeneous Poisson processes in $\mathbb{R}^d$, with emphasis on the cases where exact rates have been obtained. We also cite closely related results for the increments of the empirical process. Very recent results on the *L-increments* of a Poisson process are given, where an L-increment is the sum of points in L disjoint sets. Finally, we point out possible generalizations of the results and directions of future research.

1. Introduction and notations

Suppose we are given a random collection of points in $\mathbb{R}^d$ that follow a homogeneous Poisson process on $\mathbb{R}^d$ with parameter ψ, i.e., if for every Borel subset A of $\mathbb{R}^d$, $\eta(A)$ denotes the number of points contained in A and $\lambda(A)$ its Lebesgue measure, then $\eta(A)$ has a Poisson distribution with parameter $\psi\lambda(A)$, and the numbers of points in disjoint sets are jointly independent.

In this paper, we shall survey results on the asymptotic behaviour as $T \to \to \infty$ of the maximal and minimal number of points in certain families of sets of volume V_T contained in $H_T = [0,T]^d$, where $0 < V_T \leqq T^d$ and $\lim_{T\to\infty} V_T/T^d = = 0$. Let $\mathfrak{E}$ be a family of Borel measurable subsets of $[0,1]^d$, and let the maximal and minimal numbers be defined as

$$\Delta_T^+(V) = \max\{\eta(E) : E \in T\mathfrak{E}, \lambda(E) = V\}$$
$$\Delta_T^-(V) = \min\{\eta(E) : E \in T\mathfrak{E}, \lambda(E) = V\}.$$

(If $t \in \mathbb{R}$, and $E \subseteqq \mathbb{R}^d$, then $tE = \{tx : x \in E\}$; similarly, $t\mathfrak{E} = \{tE : E \in \mathfrak{E}\}$.)

To investigate the behaviour of L-increments we also define

$$\Delta_T^+(V,L) = \max\{\eta(E_1 \cup \cdots \cup E_L) : E_i \in T\mathfrak{E}, \lambda(E_i) = V, \{E_i\} \text{ disjoint }\},$$
$$\Delta_T^-(V,L) = \min\{\eta(E_1 \cup \cdots \cup E_L) : E_i \in T\mathfrak{E}, \lambda(E_i) = V, \{E_i\} \text{ disjoint }\},$$

1991 *Mathematics Subject Classification*. Primary 60G55; Secondary 60F15.

Key words and phrases. Homogeneous d-dimensional Poisson process, empirical process, maximal and minimal increments, L-increments, strong laws, asymptotically quasi-deterministic.

0081-6906/96/\$ 5.00 ©1996 *Akadémiai Kiadó, Budapest*

which denote the maximal and minimal number of points of η in subsets of $[0,T]^d$ given by unions of L disjoint sets E_i from $T\mathfrak{E}$, such that all sets E_i have volume V[1]. (In what follows, a family $\{E_i\}$ of sets is called "disjoint" if $\lambda(E_i \cap E_j) = 0$ for all $i \neq j$.)

In the one-dimensional case, the natural choice for the family $\mathfrak{E}$ is the set $\mathfrak{E}_{\mathrm{INT}}$ of all intervals in $[0,1]$. Multi-dimensional generalizations are e.g. $\mathfrak{E}^d_{\mathrm{CUB}}$, the set of axis-parallel cubes in $[0,1]^d$, $\mathfrak{E}^d_{\mathrm{SPH}}$, the set of spheres in $[0,1]^d$, $\mathfrak{E}^d_{\mathrm{REC}}$, the set of axis-parallel (generalized) rectangles in $[0,1]^d$, $\mathfrak{E}^d_{\rho\mathrm{CUB}}$, the set of cubes in general position in $[0,1]^d$, and $\mathfrak{E}^d_{\rho\mathrm{REC}}$, the set of rectangles in general position in $[0,1]^d$.

For $u \in \mathbb{R}$, let

$$h(u) = \begin{cases} u \log u - u + 1, & u \geqq 0, \\ \infty & u < 0, \end{cases}$$

be the familiar Chernoff function of the Poisson distribution, and let for $\phi \geqq 0$

$$u^+(\phi) = \inf\{u \geqq 1 : h(u) \geqq \phi\},$$
$$u^-(\phi) = \sup\{u \leqq 1 : h(u) \geqq \phi\}.$$

Finally, for all $x \in \mathbb{R}$, $\lfloor x \rfloor$ and $\lceil x \rceil$ denote the largest integer $\leqq x$ and the smallest integer $\geqq x$, respectively.

2. First results

Révész [14] was first to obtain the exact rate of $\Delta_T^+(1)$ by proving the following theorem.

THEOREM 2.1. *For the family of intervals $\mathfrak{E}_{\mathrm{INT}}$ and all $\varepsilon > 0$,*

$$\lfloor \psi u^+(\psi^{-1} \log T) - 1/2 - \varepsilon \rfloor \leqq \Delta_T^+(1) < \psi u^+(\psi^{-1} \log T) + 3/2 + \varepsilon$$

for all sufficiently large T with probability 1.

The proof is by direct calculation of good estimates of the relevant probabilities and application of standard Borel–Cantelli techniques.

[1] Notice that $\Delta_T^+(V,L)$ and $\Delta_T^-(V,L)$ are not necessarily measurable as functions from the underlying probability space Ω to $\mathbb{R}$. This problem is circumvented by using outer probability, denoted as $\mathbb{P}^*$, for inequalities with these quantities. In particular, if A_T is a sequence of (not necessarily measurable) subsets of Ω, we write $\lim_{T\to\infty} \mathbb{P}^*(A_T) = 1$ iff $\lim_{T\to\infty} \mathbb{P}^*(\Omega \setminus A_T) = 0$, and if $A \subset \Omega$, we say that A a.s. (almost surely, or with probability one) if there exists $\Omega_0 \subset A$ with $\mathbb{P}(\Omega_0) = 1$.

Thus the maximal number of points contained in some interval of length 1 in $[0,T]$ is one of the four values $f(T)-1$, $f(T)$, $f(T)+1$, $f(T)+2$, where $f(T)=\lfloor \psi u^+(\psi^{-1}\log T)\rfloor \sim (\log T)/(\log\log T)$.

In [14], Révész introduced the terms asymptotically quasi-deterministic and asymptotically deterministic.

DEFINITION 2.2. Let Y_T be a random process and suppose that there exist *deterministic* functions $\beta_1(T)$, $\beta_2(T)$ such that

$$\beta_1(T)\leqq Y_T\leqq \beta_2(T)$$

for all sufficiently large T with probability 1. If $\beta_2(T)-\beta_1(T)=O(1)$ as $T\to\infty$, Y_T is *asymptotically quasi-deterministic (AQD)*; if $\beta_2(T)-\beta_1(T)=$ $=o(1)$ as $T\to\infty$, it is *asymptotically deterministic (AD)*.

Clearly, Theorem 2.1 implies that $\Delta_T^+(1)$ is AQD.

2.1. Generalization to the multi-dimensional case. A straightforward generalization of Theorem 2.1 was given by Auer, Hornik and Révész [3].

THEOREM 2.3. *Let V_T be eventually monotone and $V_T=o(\log T)$ for $T\to\infty$. Then for the family of axis-parallel cubes $\mathfrak{C}^d_{\mathrm{CUB}}$ and any $\varepsilon>0$,*

$$\begin{aligned}&\left\lfloor V_T\psi u^+\left(\frac{\log(T^d/V_T)}{V_T\psi}\right)-(1/2+\varepsilon)\frac{\log\log T}{\log((\log T^d)/V_T)}\right\rfloor\\&\leqq \Delta_T^+(V_T)\\&< V_T\psi u^+\left(\frac{\log(T^d/V_T)}{V_T\psi}\right)+(\gamma+\varepsilon)\frac{\log\log T}{\log((\log T^d)/V_T)}\end{aligned}$$

for all sufficiently large T with probability 1, *where*

$$\gamma=\begin{cases} d+1/2 & \textit{if } V_T \textit{ is nondecreasing},\\ d+3/2 & \textit{if } V_T \textit{ is decreasing}.\end{cases}$$

The proof follows the line of the proof of Theorem 2.1.

Observe that Theorem 2.1 is a corollary of Theorem 2.3. Furthermore, note that Theorem 2.3 holds for quite general V_T below the *Erdős–Rényi range*[2] ($V_T\sim\log T$).

Putting some more restriction on V_T, we obtain the following

[2] We refer to Deheuvels [8], Borovkov [6], and references therein for details concerning the original Erdős–Rényi [12] law and its extensions. The Erdős–Rényi range corresponds to when $V=V_T$ is such that $V_T/\log T\to C$ with $0<C<\infty$, and we will say that the increments are above (resp. below) the Erdős–Rényi range when $C=\infty$ (resp. $C=0$).

COROLLARY 2.4. *If V_T is eventually monotone and for some $\alpha < 1$, $V_T = O((\log T)^\alpha)$ for $T \to \infty$, then $\Delta_T^+(V_T)$ is AQD for the family $\mathfrak{C}^d_{\mathrm{CUB}}$.*

COROLLARY 2.5. *If $V_T = T^{-\beta d}$ for $\beta > 0$, then for $\mathfrak{C}^d_{\mathrm{CUB}}$ and all $\varepsilon > 0$*

$$\lfloor 1 + 1/\beta - \varepsilon \rfloor \leqq \Delta_T^+(V_T) < 1 + 1/\beta + \varepsilon$$

for all sufficiently large T with probability 1. Hence in this case, $\Delta_T^+(V_T)$ is AD if β is not integer.

2.2. Generalization to volumes in and above the Erdős–Rényi range. In this section, we review further results on the increments of the one-dimensional Poisson process. In fact, these hold for general renewal processes. Nevertheless we restrict ourselves to the Poisson process to keep the presentation as simple as possible.

The following result for volumes in the Erdős–Rényi range was obtained by Bacro, Deheuvels and Steinebach [5]. Observe that the results are for the one-dimensional case only, but give exact bounds on the limiting behaviour of $\Delta_T^+(V_T)$ and $\Delta_T^-(V_T)$.

THEOREM 2.6. *Let $V_T = C \log T$, $C > 0$. Then for the family $\mathfrak{C}_{\mathrm{INT}}$*

$$\limsup_{T\to\infty} \frac{\Delta_T^+(V_T) - \psi V_T u^+(1/(C\psi))}{\log\log T} = +\frac{1}{2\log u^+(1/(C\psi))} \quad a.s.,$$
$$\liminf_{T\to\infty} \frac{\Delta_T^+(V_T) - \psi V_T u^+(1/(C\psi))}{\log\log T} = -\frac{1}{2\log u^+(1/(C\psi))} \quad a.s..$$

If furthermore $C > 1/\psi$, then

$$\limsup_{T\to\infty} \frac{\Delta_T^-(V_T) - \psi V_T u^-(1/(C\psi))}{\log\log T} = +\frac{1}{2\log u^-(1/(C\psi))} \quad a.s.,$$
$$\liminf_{T\to\infty} \frac{\Delta_T^-(V_T) - \psi V_T u^-(1/(C\psi))}{\log\log T} = -\frac{1}{2\log u^-(1/(C\psi))} \quad a.s..$$

The proof is based on a duality argument comparing the increments of a renewal process with suitable increments of the corresponding partial sum sequence. This argument was also used by Bacro [4] for related problems.

For volumes above the Erdős–Rényi range, we have the following result by Deheuvels and Steinebach [11]. Note that only the one-dimensional case is covered, and that some regularity conditions on V_T are introduced to obtain the results which give exact bounds on the limiting behaviour of $\Delta_T^+(V_T)$ and $\Delta_T^-(V_T)$.

THEOREM 2.7. *Let $\dot{V}_T = dV_T/dT$ be continuous in T. Furthermore let $V_T/\log T \to \infty$, $\log\log(T/V_T)/\log\log T \to 1$, and $\dot{V}_T/V_T = O(1/(T\log T))$ as $T \to \infty$. If $V_T(\log\log T)^2/(\log T)^3 \to \infty$ or $V_T/(\log T)^p \to 0$ for some $p > 1$ as $T \to \infty$, then for the family $\mathfrak{C}_{\mathrm{INT}}$*

$$\limsup_{T\to\infty} \frac{(2\log(T/V_T))^{1/2}}{(V_T\psi)^{1/2}\log\log(T/V_T)} \times \left(\Delta_T^+(V_T) - V_T\psi u^+\left(\frac{\log(T/V_T)}{V_T\psi}\right)\right) = 3/2 \quad a.s.,$$

$$\liminf_{T\to\infty} \frac{(2\log(T/V_T))^{1/2}}{(V_T\psi)^{1/2}\log\log(T/V_T)} \times \left(\Delta_T^+(V_T) - V_T\psi u^+\left(\frac{\log(T/V_T)}{V_T\psi}\right)\right) = 1/2 \quad a.s.,$$

$$\limsup_{T\to\infty} \frac{(2\log(T/V_T))^{1/2}}{(V_T\psi)^{1/2}\log\log(T/V_T)} \times \left(\Delta_T^-(V_T) - V_T\psi u^-\left(\frac{\log(T/V_T)}{V_T\psi}\right)\right) = -1/2 \quad a.s.,$$

$$\liminf_{T\to\infty} \frac{(2\log(T/V_T))^{1/2}}{(V_T\psi)^{1/2}\log\log(T/V_T)} \times \left(\Delta_T^-(V_T) - V_T\psi u^-\left(\frac{\log(T/V_T)}{V_T\psi}\right)\right) = -3/2 \quad a.s..$$

Again, the proof is by a duality argument which relates the increments of a renewal process to the increments of the corresponding partial sum sequence. For the case $V_T(\log\log T)^2/(\log T)^3 \to \infty$ the result is derived from the limiting behaviour of the Wiener process and a strong invariance principle due to Mason and van Zwet [13]. For $V_T/(\log T)^p \to 0$ the proof uses a theorem on increments of partial sums by Deheuvels and Steinebach [10].

3. More general families of Borel sets

Until now, we have considered the number of points of a Poisson process in intervals and cubes only. To obtain strong results for very general families of Borel sets, one has to put some conditions on these families to ensure that the maximal and minimal number of points $\Delta_T^+(V)$ and $\Delta_T^-(V)$ exhibit the "natural" behaviour in their arguments T and V. Such conditions have been introduced by Deheuvels, Einmahl, Mason and Ruymgaart [9] who analyzed the behaviour of the (increments of the) empirical process. Since the Poisson process and the empirical process are closely related, those definitions are also very well suited for our purpose.

The theorems we will present in the sequel will hold for families $\mathfrak{E}$ satisfying some of the following conditions.

(U) There exists $a_U \in (0,1)$ such that for all $a \in (0, a_U)$: whenever $E \in \mathfrak{E}$ with $\lambda(E) < a$ there exists $E' \in \mathfrak{E}$ with $\lambda(E') = a$ and $E \subset E'$.

(D) There exists $a_D \in (0,1)$ such that for all $a \in (0, a_D)$: whenever $E \in \mathfrak{E}$ with $\lambda(E) > a$ there exists $E' \in \mathfrak{E}$ with $\lambda(E') = a$ and $E' \subset E$.

(S) If $E \in \mathfrak{E}$, then $tE \in \mathfrak{E}$ for all $0 < t < 1$.

Condition (S) guarantees that $\Delta_T^+(V)$ is nondecreasing in T and $\Delta_T^-(V)$ is nonincreasing in T; condition (U) guarantees that $\Delta_T^+(V)$ is nondecreasing in V on $(0, a_U T^d)$, and condition (D) guarantees that $\Delta_T^-(V)$ is nondecreasing in V on $(0, a_D T^d)$.

Furthermore, we need to quantify how "rich" a family $\mathfrak{E}$ is. Again following [9], we introduce the following "covering" numbers. For $0 < a < 1$, let $\mathfrak{E}(a) = \{E \in \mathfrak{E} : \lambda(E) = a\}$ and set

$$K_{\mathfrak{E}}(a) = \max\left\{ m \geqq 1 : \begin{array}{l} \text{there are sets } E_1, \ldots, E_m \in \mathfrak{E}(a) \text{ such that} \\ \lambda(E_i \cap E_j) = 0 \text{ for } 1 \leqq i \neq j \leqq m \end{array} \right\},$$

for $0 < a < a(1+\nu) < 1$ write

$$M_{\mathfrak{E}}(a,\nu) = \min\left\{ m \geqq 1 : \begin{array}{l} \text{there exist Borel sets } B_1, \ldots, B_m \text{ such} \\ \text{that for any } E \in \mathfrak{E}(a) : E \subset B_i \text{ and} \\ \lambda(B_i \setminus E) \leqq \nu a \text{ for some } 1 \leqq i \leqq m \end{array} \right\},$$

for $0 < (1-\nu)a < a < 1$ write

$$N_{\mathfrak{E}}(a,\nu) = \min\left\{ m \geqq 1 : \begin{array}{l} \text{there exist Borel sets } B_1, \ldots, B_m \text{ such} \\ \text{that for any } E \in \mathfrak{E}(a) : B_i \subset E \text{ and} \\ \lambda(E \setminus B_i) \leqq \nu a \text{ for some } 1 \leqq i \leqq m \end{array} \right\},$$

(we adopt the usual conventions $\min \emptyset = \infty$ and $\max \emptyset = 0$) and let

$$K_{\mathfrak{E}} := \liminf_{a \to 0+} \frac{\log K_{\mathfrak{E}}(a)}{\log(1/a)}$$

$$M_{\mathfrak{E}} := \liminf_{\nu \to 0+} \limsup_{a \to 0+} \frac{\log M_{\mathfrak{E}}(a,\nu)}{\log(1/a)}$$

$$N_{\mathfrak{E}} := \liminf_{\nu \to 0+} \limsup_{a \to 0+} \frac{\log N_{\mathfrak{E}}(a,\nu)}{\log(1/a)}.$$

It is trivial that $aK_{\mathfrak{E}}(a) \leqq 1$ and $K_{\mathfrak{E}}(a) \leqq N_{\mathfrak{E}}(a,\nu)$ for all $0<\nu<1$, and it is readily seen that $K_{\mathfrak{E}}(a) \leqq M_{\mathfrak{E}}(a,\nu)$ provided that $0<\nu<\min(1,a^{-1}-1)$; hence, $K_{\mathfrak{E}} \leqq \min(1, M_{\mathfrak{E}}, N_{\mathfrak{E}})$.

For the families of spheres $\mathfrak{E}^d_{\mathrm{SPH}}$ and cubes $\mathfrak{E}^d_{\mathrm{CUB}}$, Deheuvels, Einmahl, Mason and Ruymgaart [9] have shown that for suitable finite and positive constants C_K and C_U which only depend on d,

$$K_{\mathfrak{E}}(a) \geqq C_K a^{-1}, \quad M_{\mathfrak{E}}(a,\nu), N_{\mathfrak{E}}(a,\nu) \leqq C_U a^{-1}\nu^{-d}$$

for all sufficiently small a and ν. Hence in this case, $K_{\mathfrak{E}} = M_{\mathfrak{E}} = N_{\mathfrak{E}} = 1$. The same holds for the family $\mathfrak{E}^d_{\rho\mathrm{CUB}}$ of cubes in general position. For the family $\mathfrak{E}^d_{\mathrm{REC}}$ of axis-parallel rectangles, Deheuvels, Einmahl, Mason and Ruymgaart [9] proved

$$M_{\mathfrak{E}}(a,\nu), N_{\mathfrak{E}}(a,\nu) \leqq C_U a^{-1}(\log(1/a))^{d-1}\nu^{1-2d},$$

which gives $K_{\mathfrak{E}} = M_{\mathfrak{E}} = N_{\mathfrak{E}} = 1$, too, whereas for the family $\mathfrak{E}^d_{\rho\mathrm{REC}}$ of rectangles in general position, Auer and Hornik [2] found

$$K_{\mathfrak{E}} = 1, \quad d \leqq M_{\mathfrak{E}}, N_{\mathfrak{E}} \leqq 1 + d(d-1)/2.$$

More general examples for $\mathfrak{E}$ and the covering numbers associated with it can be found in [2].

3.1. Results for the empirical process. Since the Poisson process is closely related to the empirical process, we cite some results for the empirical process for completeness.

Let Θ_n be the empirical process on $[0,1]^d$ generated by n uniformly distributed points. For any family $\mathfrak{E}$ of Borel sets we define the maximal and minimal number of points of the empirical process in sets of volume a,

$$\tilde{\Delta}_n^+(a) = \max\{\Theta_n(E) : E \in \mathfrak{E}, \lambda(E) = a\},$$
$$\tilde{\Delta}_n^-(a) = \min\{\Theta_n(E) : E \in \mathfrak{E}, \lambda(E) = a\}.$$

Deheuvels, Einmahl, Mason and Ruymgaart [9] gave the following results on the limiting behaviour of $\tilde{\Delta}_n^+(a_n)$ and $\tilde{\Delta}_n^-(a_n)$.

THEOREM 3.1. *Let $\mathfrak{E}$ be a family of Borel sets such that conditions* (U) *and* (D) *are satisfied and $K_{\mathfrak{E}} = M_{\mathfrak{E}} = N_{\mathfrak{E}} = 1$.*
If $na_n = C\log n$, then

$$\lim_{n\to\infty} \frac{\tilde{\Delta}_n^+(a_n)}{\log n} = Cu^+(1/C) \quad a.s.. \tag{1}$$

If $na_n = C\log n$, with $C \geqq 1$, then

$$\lim_{n\to\infty} \frac{\tilde{\Delta}_n^-(a_n)}{\log n} = Cu^-(1/C) \quad a.s.. \tag{2}$$

If $na_n = C \log n$, *with* $C < 1$, *then*

$$\lim_{n\to\infty} \tilde{\Delta}_n^-(a_n) = 0 \quad a.s.. \tag{3}$$

If $na_n / \log n \searrow 0$ *and* $\log((\log n)/na_n)/\log n \searrow 0$ *for* $n \to \infty$, *then*

$$\lim_{n\to\infty} \frac{\log((\log n)/na_n)}{\log n} \tilde{\Delta}_n^+(a_n) = 1 \quad a.s.. \tag{4}$$

If $a_n = n^{-1-\beta}$, *where* $\beta > 0$ *and* $1/\beta$ *is not an integer, then*

$$\lim_{n\to\infty} \tilde{\Delta}_n^+(a_n) = \lfloor 1 + 1/\beta \rfloor \quad a.s.. \tag{5}$$

It is well-known that for any Borel set $E \in [0,1]^d$, the distribution of $\Theta_n(E)$ is equal to the distribution of $\eta(TE)$ given that $\eta([0,T]^d) = n$, where η is a Poisson process with parameter $\psi = 1$. Thus, one would expect that $\tilde{\Delta}_n^+(a) \approx \Delta_T^+(V)$ and $\tilde{\Delta}_n^-(a) \approx \Delta_T^-(V)$ for $n = T^d$ and $a = V/T^d$. And in fact, (1) and (2) exactly correspond to Theorem 2.6, (4) corresponds to Theorem 2.3, and (5) corresponds to Corollary 2.5.

3.2. Results for the Poisson process. The following results are due to Auer and Hornik [2] . We start with a quite general theorem on $\Delta_T^+(V_T)$; observe the very mild restrictions on V_T, only assuming that V_T is eventually monotone.

THEOREM 3.2. *Let* $\mathfrak{E}$ *be a family of Borel sets such that conditions* (U), (D), *and* (S) *are satisfied and* $0 < K_{\mathfrak{E}} \leqq M_{\mathfrak{E}} < \infty$. *If* V_T *is eventually monotone and* $V_T/T^d \to 0$ *as* $T \to \infty$, *then for all* $\varepsilon > 0$

$$\begin{aligned} &\left\lfloor (1-\varepsilon) V_T \psi u^+ \left(K_{\mathfrak{E}} \frac{\log(T^d/V_T)}{V_T \psi} \right) \right\rfloor \\ &\leqq \Delta_T^+(V_T) \\ &< (1+\varepsilon) V_T \psi u^+ \left(M_{\mathfrak{E}} \frac{\log(T^d/V_T)}{V_T \psi} \right), \end{aligned}$$

$$\begin{aligned} &V_T \psi \left(u^- \left(N_{\mathfrak{E}} \frac{\log(T^d/V_T)}{V_T \psi} \right) - \varepsilon \right) \\ &< \Delta_T^-(V_T) \\ &\leqq \left\lceil V_T \psi \left(u^- \left(K_{\mathfrak{E}} \frac{\log(T^d/V_T)}{V_T \psi} \right) + \varepsilon \right) \right\rceil \end{aligned}$$

for all sufficiently large T with probability 1.

Idea of the Proof. The proof starts as usual by calculating appropriate bounds on certain probabilities. The really interesting part is how strong results can be obtained for rather arbitrary V_T which possibly exhibit strong variation in growth and even need not to be continuous. This is achieved by sophisticatedly choosing points T_n such that the theorem holds for T_n, n sufficiently large (by an application of the Borel–Cantelli lemma), *and* the variation of V_T is not too big in the interval (T_{n-1}, T_n) such that the gap between T_{n-1} and T_n can be closed using the regularity conditions on $\mathfrak{E}$ and V_T. As a simplified example where V_T is nondecreasing and continuous and $K_{\mathfrak{E}} = M_{\mathfrak{E}} = 1$, one can choose

$$T_n = \sup\{T > T_{n-1} : m(T) \leqq m(T_{n-1}) + n^{-1/2}\},$$

where $m(T) = V_T\psi u^+(\log(T^d/V_T)/V_T\psi)$. For details, see [2].

Note that for the special cases where $\mathfrak{E} = \mathfrak{E}_{\mathrm{INT}}$ (hence $d = 1$) and much more restrictions on V_T are imposed, Theorem 3.2 is weaker than e.g. Theorem 2.7, because it does not provide any remainder terms. In fact, for $d > 1$, such terms have thus far only been obtained for volumes not above the Erdős–Rényi range (see [2]).

THEOREM 3.3. *Let $\mathfrak{E}$ be a family of Borel sets such that conditions* (U), (D), *and* (S) *are satisfied,* $0 < K_{\mathfrak{E}} \leqq M_{\mathfrak{E}} < \infty$ *and*

$$\begin{aligned} K_{\mathfrak{E}}(a) &\geqq C_K a^{-K_{\mathfrak{E}}} \\ M_{\mathfrak{E}}(a,\nu) &\leqq C_M a^{-M_{\mathfrak{E}}}(\log 1/a)^{\alpha}\nu^{-\beta} \end{aligned}$$

for some finite and positive C_K and C_M. If V_T is eventually monotone and $V_T = O(\log T)$ as $T \to \infty$, then for all $\varepsilon > 0$

$$\begin{aligned} &\left\lfloor V_T\psi u^+\left(K_{\mathfrak{E}}\frac{\log(T^d/V_T)}{V_T\psi}\right) - (1/2+\varepsilon)\frac{\log\log(T^d/V_T)}{\log u^+\left(K_{\mathfrak{E}}\frac{\log(T^d/V_T)}{V_T\psi}\right)}\right\rfloor \\ &\leqq \Delta_T^+(V_T) \\ &< V_T\psi u^+\left(M_{\mathfrak{E}}\frac{\log(T^d/V_T)}{V_T\psi}\right) + (\gamma+\varepsilon)\frac{\log\log(T^d/V_T)}{\log u^+\left(M_{\mathfrak{E}}\frac{\log(T^d/V_T)}{V_T\psi}\right)} \end{aligned}$$

for all sufficiently large T with probability 1, *where*

$$\gamma = \begin{cases} \alpha+\beta+1/2 & \textit{if } V_T \textit{ is nondecreasing,} \\ \alpha+\beta+3/2 & \textit{if } V_T \textit{ is decreasing.} \end{cases}$$

If $\limsup_{T\to\infty} V_T\psi/\log T^d < K_{\mathfrak{C}}$, *then*

$$\lim_{T\to\infty} \Delta_T^-(V_T)=0 \quad a.s..$$

REMARK 3.4. For $\mathfrak{C}^d_{\mathrm{CUB}}$, a sharper bound for the largest V_T such that $\Delta_T^-(V_T)=0$ (i.e., on the volume of the largest empty cube) was proved by Deheuvels [7].

From the bounds on the covering numbers given before Section 3.1, we find that in the above theorem, we can take $\alpha=0$ and $\beta=d$ for $\mathfrak{C}^d_{\mathrm{CUB}}$ and $\mathfrak{C}^d_{\mathrm{SPH}}$, $\alpha=d-1$ and $\beta=2d-1$ for $\mathfrak{C}^d_{\mathrm{REC}}$, $\alpha=0$ and $\beta=d(d+1)/2$ for $\mathfrak{C}^d_{\rho\mathrm{CUB}}$, and $\alpha=d-1$ and $\beta=d(d+1)/2$ for $\mathfrak{C}^d_{\rho\mathrm{REC}}$. Hence, comparing Theorem 3.3 with the results of Section 2, we find that Theorem 2.3 is a corollary of Theorem 3.3. Comparing with Theorem 2.6, we have the following corollary.

COROLLARY 3.5. *If, in addition to the assumptions of Theorem 3.3,* $K_{\mathfrak{C}}=M_{\mathfrak{C}}=1$ *and* $V_T=C\log T^d$, *then*

$$\limsup_{T\to\infty} \frac{\Delta_T^+(V_T)-\psi V_T u^+(1/(C\psi))}{\log\log T} \leqq \frac{\alpha+\beta+1/2}{\log u^+(1/(C\psi))} \quad a.s.$$

$$\liminf_{T\to\infty} \frac{\Delta_T^+(V_T)-\psi V_T u^+(1/(C\psi))}{\log\log T} \geqq \frac{-1/2}{\log u^+(1/(C\psi))} \quad a.s..$$

Thus, for the special case $\mathfrak{C}=\mathfrak{C}_{\mathrm{INT}}$ where Theorem 2.6 applies, we get the same order of the remainder terms, but a worse constant in the upper bound.

4. On L-increments

Theorem 3.2 implies that for $V_T/(\log T)\to\infty$, $\lim_{T\to\infty} \Delta_T^{\pm}(V_T)/(V_T\psi)=1$ a.s. and thus, since all sets in $T\mathfrak{C}$ contain about $V_T\psi$ points,

$$\lim_{T\to\infty} \frac{\Delta_T^+(V_T,L_T)}{L_T V_T\psi} = \lim_{T\to\infty} \frac{\Delta_T^-(V_T,L_T)}{L_T V_T\psi} = 1 \quad \text{a.s.}$$

provided that $L_T \leqq K_{\mathfrak{C}}(V_T/T^d)$.

To investigate more closely the number of sets containing a certain number of points, or vice versa, the number of points in the union of a certain number of sets, we will give results on the limiting behaviour of $\Delta_T^+(V_T,L_T)$ and $\Delta_T^-(V_T,L_T)$. We only give weak, i.e. "in probability" limit theorems. Strong, i.e. "with probability one" results can be obtained analogously to Theorem 3.2 under suitable regularity conditions. The weak results already clearly demonstrate the rather surprising fact that there are many sets containing about the maximal and minimal numbers of points, respectively. A more detailed investigation can be found in [1].

THEOREM 4.1. *If $L_T/K_{\mathfrak{E}}(V_T/T^d) \to 0$ as $T \to \infty$, then for all $\nu, \varepsilon > 0$,*

$$\lim_{T\to\infty} \mathbf{P}^* \Bigg\{ L_T \left\lfloor (1-\varepsilon) V_T \psi u^+ \left(\frac{1}{V_T\psi} \log \frac{K_{\mathfrak{E}}(T^d/V_T)}{L_T} \right) \right\rfloor \leqq \Delta_T^+(V_T, L_T) < (1+\nu+\varepsilon) L_T V_T \psi u^+ \left(\frac{1}{V_T\psi} \log \frac{M_{\mathfrak{E}}(T^d/V_T,\nu)}{L_T} \right) \Bigg\} = 1,$$

$$\lim_{T\to\infty} \mathbf{P}^* \Bigg\{ (1-\nu) L_T V_T \psi \left(u^- \left(\frac{1}{V_T\psi} \log \frac{N_{\mathfrak{E}}(T^d/V_T,\nu)}{L_T} \right) - \varepsilon \right) < \Delta_T^-(V_T, L_T) \leqq L_T \left\lceil V_T \psi \left(u^- \left(\frac{1}{V_T\psi} \log \frac{K_{\mathfrak{E}}(T^d/V_T)}{L_T} \right) + \varepsilon \right) \right\rceil \Bigg\} = 1.$$

If $\limsup_{T\to\infty} (\psi V_T + \log L_T)/\log T^d < K_{\mathfrak{E}}$, *then*

$$\lim_{T\to\infty} \mathbf{P}^* \{ \Delta_T^-(V_T, L_T) = 0 \} = 1.$$

Observe that no regularity assumptions on $\mathfrak{E}$ or V_T are necessary to obtain weak laws for $\Delta_T^+(V_T, L_T)$ and $\Delta_T^-(V_T, L_T)$. To give a more intuitive picture of the situation, we restrict ourselves to two special cases.

COROLLARY 4.2. *If $K_{\mathfrak{E}} = M_{\mathfrak{E}} = N_{\mathfrak{E}} = 1$, $V_T = C \log T^d$, and $L_T = \lfloor T^{\gamma d} \rfloor$ for some $0 \leqq \gamma < 1$, then*

$$\lim_{T\to\infty} \frac{\Delta_T^{\pm}(V_T, L_T)}{L_T V_T \psi} = u^{\pm} \left(\frac{1-\gamma}{C\psi} \right) \quad \textit{i.p.}.$$

This corollary quantifies what one expects from Theorem 2.6. The next corollary is much more surprising since it states that there is a rather large number of sets of volume 1 which contain about the maximal possible number of points, and similarly a rather large number of empty sets of volume 1.

COROLLARY 4.3. *If $K_{\mathfrak{E}} = M_{\mathfrak{E}} = N_{\mathfrak{E}} = 1$, and $L_T = \lfloor T^{\gamma d} \rfloor$ for some $0 \leqq \gamma < 1$, then*

$$\lim_{T\to\infty} \frac{\Delta_T^+(1, L_T)}{L_T \Delta_T^+(1,1)} = 1 - \gamma \quad \textit{i.p.}$$

and

$$\lim_{T\to\infty} \Delta_T^-(1, L_T) = 0 \quad \textit{i.p.}.$$

A more general result for the case where $\Delta_T^{\pm}(V_T, L_T)$ is about $L_T \Delta_T^{\pm}(V_T, 1)$ is given in the following corollary.

COROLLARY 4.4. *Let* $K_{\mathfrak{C}} = M_{\mathfrak{C}} = N_{\mathfrak{C}} = 1$ *and* $\log(L_T)/\log(T^d/V_T) \to 0$ *as* $T \to \infty$. *If* $\liminf_{T\to\infty} V_T > 0$, *then*

$$\lim_{T\to\infty} \frac{\Delta_T^+(V_T, L_T)}{L_T \Delta_T^+(V_T, 1)} = 1 \quad i.p.,$$

if $\liminf_{T\to\infty} V_T\psi/\log T^d > 1$, *then*

$$\lim_{T\to\infty} \frac{\Delta_T^-(V_T, L_T)}{L_T \Delta_T^-(V_T, 1)} = 1 \quad i.p..$$

5. Conclusion and directions of research

We have presented results on the maximal and minimal increments of a homogeneous Poisson process. We started with the univariate case and gave results for intervals of length below, in, and above the Erdős–Rényi range. The results in and above the Erdős–Rényi range are sharp in the sense that exact remainder terms are known. In the multivariate case we used a very general definition of "increments" for which surprisingly general bounds can be obtained. A drawback of these results is that currently, no remainder terms are known above the Erdős–Rényi range. We also presented results on L-increments which show that surprisingly many sets contain about the maximal or minimal possible number of points.

In future research, it would be interesting to establish exact remainder terms for the multivariate case and also for L-increments. From the results on L-increments, one can obtain results on the number of sets containing a certain number of points, but the conclusions are rather weak. More precise results are desirable. Finally, the results might extend to general multivariate processes with independent increments.

REFERENCES

[1] AUER, P. and HORNIK, K., The number of points of an empirical or Poisson process covered by unions of sets (submitted for publication).

[2] AUER, P. and HORNIK, K., On the number of points of a homogeneous Poisson process, *J. Multivariate Anal.* **48** (1994), 115–156.

[3] AUER, P., HORNIK, K. and RÉVÉSZ, P., Some limit theorems for the homogeneous Poisson process, *Statist. Probab. Lett.* **12** (1991), 91–96. *MR* **92j**:60037

[4] BACRO, J.-N., Encadrements limites de lois de type Erdős–Rényi–Shepp, *Publ. Inst. Statist. Univ. Paris* **31** (1986), 1–14. *MR* **88k**:60056

[5] BACRO, J.-N., DEHEUVELS, P. and STEINEBACH, J., Exact convergence rates in Erdős–Rényi type theorems for renewal processes, *Ann. Inst. H. Poincaré Probab. Statist.* **23** (1987), 195–207. *MR* **88k**:60055

[6] BOROVKOV, K. A., The functional form of the Erdős–Rényi law of large numbers, *Teor. Veroyatnost. i Primenen.* **35** (1990), 758–762. translation in *Theory Probab. Appl.* **35** (1990), 762–766. *MR* **92e**:60060

[7] DEHEUVELS, P., On the Erdős–Rényi theorem for random fields and sequences and its relationships with the theory of runs and spacings, *Z. Wahrsch. Verw. Gebiete* **70** (1985), 91–115. *MR* **86k**:60045

[8] DEHEUVELS, P., Functional Erdős–Rényi laws, *Studia Sci. Math. Hungar.* **26** (1991), 261–295. *MR* **93i**:60056

[9] DEHEUVELS, P., EINMAHL, J. H. J., MASON, D. and RUYMGAART, F. H., The almost sure behaviour of maximal and minimal multivariate k_n-spacings, *J. Multivariate Anal.* **24** (1988), 155–176. *MR* **89a**:60078

[10] DEHEUVELS, P. and STEINEBACH, J., Exact convergence rates in strong approximation laws for large increments of partial sums, *Probab. Theory Related Fields* **76** (1987), 369–393. *MR* **88m**: 60081

[11] DEHEUVELS, P. and STEINEBACH, J., Sharp rates for increments of renewal processes, *Ann. Probab.* **17** (1989), 700–722. *MR* **90a**:60058

[12] ERDŐS, P. and RÉNYI, A., On a new law of large numbers, *J. Analyse Math.* **23** (1970), 103–111. *MR* **42** # 6907

[13] MASON, D. M. and ZWET, W. VAN, A note on the strong approximation to the renewal process, *Publ. Inst. Statist. Univ. Paris* **32** (1987), 81–91. *MR* **88j**:60066

[14] RÉVÉSZ, P., How random is random?, *Probab. Math. Statist.* **4** (1984), 109–116. *MR* **86c**:60052

(Received September 3, 1994)

INSTITUT FÜR GRUNDLAGEN DER INFORMATIK
TECHNISCHE UNIVERSITÄT GRAZ
KLOSTERWIESGASSE 32/II
A–8010 GRAZ
AUSTRIA

pauer@ igi.tu-graz.ac.at

INSTITUT FÜR STATISTIK UND WAHRSCHEINLICHKEITSTHEORIE
TECHNISCHE UNIVERSITÄT WIEN
WIEDNER HAUPTSTRASSE 8–10/1071
A–1040 WIEN
AUSTRIA

Kurt.Hornik@ tuwien.ac.at

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 15–25

TRIGONOMETRIC SERIES AND UNIFORM DISTRIBUTION MOD 1

I. BERKES* and W. PHILIPP

Dedicated to Pál Révész on his 60th birthday

Abstract

We give a survey on recent results on estimates of the discrepancy of sequences $\{n_k\omega\}$ mod 1 and their connection with trigonometric series.

1. Some facts from the theory of uniform distribution mod 1

Let $\{\eta_1, \dots, \eta_N\}$ be a finite sequence of real numbers with $0 \leqq \eta_k < 1$. Write

(1.1) $$F_N(s) := N^{-1}\text{card}\{k \leqq N : \eta_k \leqq s\}, \quad 0 \leqq s < 1$$

and

(1.2) $$D_N := \sup_{0 \leqq s < 1} |F_N(s) - s|.$$

Recall that D_N is the discrepancy of $\{\eta_1, \dots, \eta_N\}$. An infinite sequence $\{\eta_k,\ k \geqq 1\}$ is called uniformly distributed if $D_N \to 0$ as $N \to \infty$. Here D_N denotes discrepancy of the initial segment $\{\eta_1, \dots, \eta_N\}$.

Two inequalities link the discrepancy to exponential sums. The first one, Koksma's inequality, holds for functions f of bounded variation $V(f)$ over $[0, 1]$. For any $N \geqq 1$

(1.3) $$\left| \frac{1}{N} \sum_{k \leqq N} f(\eta_k) - \int_0^1 f(x)dx \right| \leqq V(f) D_N,$$

in particular,

(1.4) $$\frac{1}{N} \left| \sum_{k \leqq N} e^{2\pi i \eta_k} \right| \leqq 4 D_N.$$

1991 *Mathematics Subject Classification.* Primary 11K38.

Key words and phrases. Trigonometric and Walsh series, uniform distribution mod 1, discrepancy.

* Supported by the Hungarian National Foundation for Scientific Research Grant No. 1905.

The second one, the Erdős–Turán inequality says that for any $m \geqq 1$

$$D_N \leqq \frac{6}{m+1} + \frac{4}{\pi} \sum_{h=1}^{m} \frac{1}{h} \cdot \frac{1}{N} \left| \sum_{k \leqq N} e^{2\pi i h \eta_k} \right|. \tag{1.5}$$

The inequalities (1.4) and (1.5) can be considered a quantitative version of the Weyl criterion: $D_N \to 0$ iff

$$\frac{1}{N} \sum_{k \leqq N} e^{2\pi i h \eta_k} \to 0 \qquad \text{for all} \quad h = 1, 2, \ldots .$$

2. Sequences $\{n_k \omega\}$ mod 1 (general case)

In what follows, $\{n_k,\ k \geqq 1\}$ will denote an increasing sequence of positive integers and $([0,1), \mathcal{B}, \mathbf{P})$ will denote the unit interval with Lebesgue measurability and Lebesgue measure. We introduce

$$\eta_k = \eta_k(\omega) := n_k \omega \quad \mod 1, \tag{2.1}$$

and consider them as random variables on the probability space $([0,1), \mathcal{B}, \mathbf{P})$. This definition reveals the probabilistic nature of the notions introduced above. Indeed, $F_N(s) = F_N(s, \omega)$ defined in (1.1) becomes the empirical distribution function of the sequence $\{\eta_k,\ k \geqq 1\}$ at stage N and $D_N = D_N(\omega)$ becomes the Kolmogorov–Smirnov statistic. The random variables η_k, although in general dependent, are identically distributed with the uniform distribution (in the probabilistic sense) as their common distribution. Indeed, it is easy to see that

$$\mathbf{P}\{\omega :\ \eta_k(\omega) \leqq x\} = x \qquad \text{for } 0 \leqq x \leqq 1,\ k \geqq 1.$$

In this context the inequalities (1.4) and (1.5) establish a link with trigonometric series, a connection frequently exploited.

The question of the asymptotic behaviour of $D_N(\omega)$ has a long and illustrious history. Weyl [28] proved in his seminal paper that

$$D_N(\omega) \to 0 \qquad \text{a.s.}$$

for any $\{n_k\}$. This was later improved independently by Cassels [8] and Erdős and Koksma [13] to

$$D_N(\omega) \ll N^{-\frac{1}{2}} (\log N)^{\frac{5}{2}+\varepsilon} \qquad \text{a.s.} \tag{2.2}$$

for any $\varepsilon > 0$. The exponent $\frac{5}{2} = \frac{3}{2} + 1$ in the logarithmic factor stems from two sources. The exponent $\frac{3}{2}$ is typical for results on the growth of partial sums of orthogonal functions. Since only the second order properties of η_k were used in the papers by Cassels and Erdős and Koksma, the exponent $\frac{3}{2}$ fits this pattern. The remaining exponent 1 represents the "usual defect" of the Erdős–Turán inequality.

It took some 30 years to improve on (2.2): R. C. Baker [1] was able to reduce the exponent $\frac{5}{2}$ to $\frac{3}{2}$. He used R. Hunt's [14] deep maximal inequality for trigonometric series. This accounts for the reduction of the exponent $\frac{3}{2}$, typical for orthogonal functions, to $\frac{1}{2}$. Since Baker uses the Erdős–Turán inequality, the "usual defect" of 1 still remains. He conjectured that

$$D_N(\omega) \ll N^{-\frac{1}{2}}(\log N)^{\varepsilon} \qquad \text{a.s.} \tag{2.3}$$

for any $\varepsilon > 0$. Earlier, in his Nijenrode Lecture, Erdős [11] had conjectured that

$$D_N(\omega) = O(N^{-\frac{1}{2}}(\log\log N)^c) \qquad \text{a.s.} \tag{2.4}$$

for some $c(\geqq \frac{1}{2})$.

Both conjectures were disproved by Berkes and Philipp [6].

THEOREM A. *Let f be a non-decreasing function with*

$$\sup_{k \geqq 1} \frac{f(k^2)}{f(k)} < \infty \tag{2.5}$$

and

$$\sum_{k=1}^{\infty} \frac{1}{k(f(k))^2} = \infty.$$

Then there exists an increasing sequence $\{n_k\}$ of integers such that

$$\limsup_{N\to\infty} \frac{\left| \sum_{k \leqq N} \cos 2\pi n_k \omega \right|}{\sqrt{N} f(N)} = \infty \qquad \text{a.s.}.$$

Theorem A with $f(k) = (\log k)^{1/2}$, combined with Koksma's inequality (1.4) shows that for some increasing sequence $\{n_k\}$ of integers we have

$$\limsup_{N\to\infty} (N/\log N)^{\frac{1}{2}} D_N(\omega) = \infty \qquad \text{a.s.} \tag{2.6}$$

and thus (2.3) and (2.4) fail.

We do not know what the optimal exponent of $\log N$ in (2.2) is and in view of the very different arguments used in Baker's upper bound and our lower bound (2.6), it is hard to make a reasonable conjecture. Certainly, (2.6) cannot be improved by improving Theorem A above. Indeed, if f is a non-decreasing function with

$$\sum_{k=1}^{\infty} \frac{1}{k(f(k))^2} < \infty$$

then by the Carleson [7] convergence theorem the series

$$\sum_{k=1}^{\infty} \frac{1}{k^{1/2} f(k)} \cos 2\pi n_k \omega$$

converges a.e. and thus by the Kronecker lemma

$$\lim_{N\to\infty} \frac{\sum_{k\leqq N} \cos 2\pi n_k \omega}{\sqrt{N} f(N)} = 0 \qquad \text{a.e..}$$

Hence the lower bound given by Theorem A is sharp. Thus to improve (2.6) one would need a result of the type of Theorem A for a function of bounded variation different from the cosine or sine function.

REMARK 1. Let $\{w_n\}$ be the Walsh system. Révész asked what are the functions f such that

$$\lim_{N\to\infty} \frac{\sum_{k\leqq N} w_{n_k}(x)}{\sqrt{N} f(N)} = 0 \qquad \text{a.e.} \tag{2.7}$$

for all increasing sequences $\{n_k\}$ of integers. (See Révész [22] p. 117). The Walsh analogue of Theorem A, proved in Berkes and Philipp [6] answers Révész's question: under the mild regularity condition (2.5), relation (2.7) holds for all $\{n_k\}$ iff

$$\sum_{k=1}^{\infty} \frac{1}{k(f(k))^2} < \infty.$$

REMARK 2. The sequence $\{n_k\}$ in Theorem A satisfies

$$n_{k+1}/n_k \geqq 1 + 1/k^{\alpha}$$

for any $\alpha > 1/2$ and thus the same holds for the sequence $\{n_k\}$ underlying (2.6).

3. The sequence $\{k\omega\}$ mod 1

It is clear from relation (1.5) that $D_N(\omega)$ will be very small almost surely since the last term in (1.5) reduces to a geometric series. For the details see Kuipers and Niederreiter [17], p. 131, Ex. 3.13. for this approach which yields

$$D_N(\omega) \ll N^{-1}(\log N)^{2+\varepsilon} \qquad \text{a.s.}. \tag{3.1}$$

However, according to a classical result of Khintchine [16]

$$D_N(\omega) \ll N^{-1} \log N \cdot g(\log\log N) \qquad \text{a.s.} \tag{3.2}$$

for any non-decreasing positive function g with

$$\sum_{n=1}^{\infty} 1/g(n) < \infty.$$

This improves the exponent $2+\varepsilon$ in (3.1) to $1+\varepsilon$.

We note that Kesten [14] proved that D_N, when normalized by $(\log N \log\log N)/N$, converges in measure to $2/\pi^2$.

4. Fast growing sequences $\{n_k\}$

It has been known since the early 1950's that for exponentially increasing sequences $\{n_k\}$, that is for sequences satisfying a Hadamard gap condition

$$n_{k+1}/n_k > 1+\rho, \quad (\rho > 0) \quad k = 1, 2, \ldots \tag{4.1}$$

the random variables η_k, defined in (2.1), behave as if they were independent. The following strong approximation theorems are, in essence, due to Philipp and Stout [21]. Let $\Omega = [0,1)^2$ be the unit square with Lebesgue measurability. Write $\omega = (\omega_1, \omega_2) \in \Omega$.

THEOREM B. *Let $\{n_k\}$ satisfy condition* (4.1). *Then there exists a sequence $\{Y_k(\omega_1, \omega_2)\}_{k=1}^{\infty}$ of independent standard Gaussian random variables defined on $[0,1)^2$ such that for almost all $\omega = (\omega_1, \omega_2) \in [0,1)^2$*

$$2^{\frac{1}{2}} \sum_{k \leqq N} \cos 2\pi n_k \omega_1 - \sum_{k \leqq N} Y_k(\omega_1, \omega_2) \ll N^{\frac{1}{2}-\lambda}$$

for some $\lambda > 0$, depending on ρ only.

By treating $e^{2\pi i \cdot} = (\cos 2\pi \cdot, \sin 2\pi \cdot)$ as a two-dimensional random vector, the same proof yields

THEOREM C. *Let $\{n_k\}$ satisfy condition* (4.1). *Then there exists a sequence $\{Y_k(\omega_1,\omega_2)\}_{k=1}^{\infty}$ of independent standard complex valued Gaussian random variables defined on $[0,1)^2$ such that for almost all $\omega=(\omega_1,\omega_2)\in$ $\in[0,1)^2$*

$$2^{\frac{1}{2}}\sum_{k\leqq N} e^{2\pi i n_k\omega_1}-\sum_{k\leqq N} Y_k(\omega_1,\omega_2)\ll N^{\frac{1}{2}-\lambda}$$

for some $\lambda>0$, depending on ρ only.

Since for almost all $\omega\in[0,1)^2$

$$\limsup_{N\to\infty}\frac{\left|\sum_{k\leqq N} Y_k(\omega_1,\omega_2)\right|}{\sqrt{2N\log\log N}}=1$$

by the classical law of the iterated logarithm for 2-dimensional standard Gaussian random vectors, Theorem C implies that for almost all $\omega_1\in[0,1)$

$$\limsup_{N\to\infty}\frac{\left|\sum_{k\leqq N} e^{2\pi i n_k\omega_1}\right|}{\sqrt{N\log\log N}}=1. \tag{4.2}$$

The upper bound in (4.2) is due to Salem and Zygmund [24] and the lower bound to Erdős and Gál [12]. Of course, Theorem B also implies the central limit theorem, due to Salem and Zygmund [23]

$$\mathbf{P}\left\{\omega:\ 2^{\frac{1}{2}}\sum_{k\leqq N}\cos 2\pi n_k\omega\leqq xN^{\frac{1}{2}}\right\}\to\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\,dt. \tag{4.3}$$

The combination of (1.4) and (4.2) immediately implies a lower bound in the law of iterated logarithm for D_N:

$$\limsup_{N\to\infty}\frac{N D_N(\omega)}{\sqrt{N\log\log N}}\geqq\frac{1}{4}\qquad\text{a.s.}. \tag{4.4}$$

An upper bound was proved by Philipp [18]:

THEOREM D. *Under* (4.1) *we have with probability* 1

$$\limsup_{N\to\infty}\frac{N D_N(\omega)}{\sqrt{N\log\log N}}\leqq C(\rho) \tag{4.5}$$

where

$$C(\rho)\ll\frac{1}{\rho}\qquad \textit{as } \rho\downarrow 0.$$

This established a long-standing conjecture of Erdős and Gál [12]. In Philipp [19] it was shown that both (4.4) and (4.5) continue to hold for not necessarily integer sequences $\{n_k\}$.

Theorems B and C continue to hold for some sequences $\{n_k\}$ growing subexponentially, namely for sequences satisfying the Erdős gap condition

$$n_{k+1}/n_k \geqq 1 + 1/k^\alpha \qquad \alpha < 1/2. \tag{4.6}$$

This was proved by Berkes [2]. Consequently, the law of the iterated logarithm (4.2) and the central limit theorem (4.3) remain also valid under (4.6). These two results were proved earlier by Takahashi [25], [26] and Erdős [10], respectively. This, of course, implies (4.4) for sequences $\{n_k\}$ satisfying the Erdős gap condition (4.6). In Philipp [18] the question was raised whether or not the upper bound (4.5) continues to hold for sequences $\{n_k\}$ satisfying the gap condition (4.6). Recently a negative answer to this question was given by Berkes and Philipp [6]:

THEOREM E. *Let $\{\rho(k) = \rho_k\}$ be a non-increasing sequence with*

$$\sup_{k \geqq 1} \frac{\rho(k)}{\rho(k^2)} < \infty.$$

Then there exists a sequence $\{n_k\}$ of integers satisfying

$$n_{k+1}/n_k \geqq 1 + \rho_k \qquad \text{for } k \geqq 1$$

such that for the discrepancy $D_N(\omega)$ of $(n_k\omega)$ mod 1 we have

$$\limsup_{N\to\infty} \frac{N D_N(\omega)}{\log\log(1/\rho_N)\sqrt{N \log\log N}} \geqq c > 0 \qquad \text{a.e.}. \tag{4.7}$$

Here c is an absolute constant.

REMARK 1. As mentioned before, under (4.6) $\{\cos 2\pi n_k x\}$ obeys the law of the iterated logarithm (see Takahashi [25], [26] and in fact the same holds for $f(n_k x)$ provided that

$$f(x+1) = f(x), \qquad \int_0^1 f(x)dx = 0, \qquad f \in \operatorname{Lip}\beta, \qquad \beta > 1/2$$

(see Dhompongsa [9]). Theorem E and its proof show that the situation is radically different for indicator functions f: in this case no gap condition weaker than (4.1) implies the LIL for $f(n_k x)$.

REMARK 2. While the lower bound $\frac{1}{4}$ in (4.4) holds for all sequences $\{n_k\}$ subject to (4.6), Theorem E shows that there are sequences $\{n_k\}$ of integers satisfying (4.1) for which the lower bound in (4.4) can be improved to $c \log\log 1/\rho$. This is an indication of the rate at which the upper bound in (4.5) breaks down as ρ gets small.

Let $\{n_k\}$ be the sequence consisting of all integers of the form $q_1^{\alpha_1} \cdots q_\tau^{\alpha_\tau}$ ($\alpha_i \geqq 0$ integers), arranged in increasing order where $\{q_1, \dots, q_\tau\}$ is a finite set of coprime integers. Let τ^* denote the total number of primes occurring in the prime factorizations of $q_1, \dots, q_\tau$. It follows from a theorem of Tijdeman [27] that $\{n_k\}$ satisfies $n_{k+1}/n_k \geqq 1 + 1/k^\alpha$ with an effectively computable $\alpha > 0$ depending on τ^*. Unfortunately, for $\tau \geqq 2$ we have $\alpha > 1/2$ and thus the above theory does not apply. Still, as is shown in Philipp [20], Theorems B, C and D continue to hold with constants λ and C depending only on τ^*. In fact Theorem D is obtained in the following slightly stronger form:

THEOREM F. *Let $\alpha < 1/(4\tau)$. There is a constant C, depending on τ^* only, with the following property. For almost all $\omega \in [0, 1)$ there is an $N_0 = N_0(\omega, \alpha)$ such that for all $N \geqq N_0$ and all s and t with $0 \leqq s < t \leqq 1$*

$$\max_{k \leqq N} k|F_k(t) - F_k(s) - (t - s)| \leqq C(t - s)^\alpha (N \log\log N)^{\frac{1}{2}} + N^{\frac{1}{2}}.$$

As an immediate consequence, we obtain the following corollary. For a proof, see [19], pp. 325–326.

COROLLARY. *Define*

$$f_N(t) = N(F_N(t) - t)(2N \log\log N)^{-\frac{1}{2}}, \qquad 0 \leqq t \leqq 1, \ N \geqq 1.$$

Then the sequence $\{f_N(t),\ N \geqq 1\}$ is with probability 1 relatively compact in $D[0, 1]$ endowed with the supremum norm.

5. Critical behaviour of trigonometric series and irregularity of $\{n_k\omega\}$ mod 1

Erdős [10] proved that $\cos 2\pi n_k \omega$ satisfies the central limit theorem (4.3) if

$$n_{k+1}/n_k \geqq 1 + c_k/\sqrt{k}, \qquad c_k \to \infty \tag{5.1}$$

and this result is sharp, i.e. for any $c > 0$ there exists a sequence $\{n_k\}$ with

$$n_{k+1}/n_k \geqq 1 + c/\sqrt{k} \qquad (k \geqq k_0) \tag{5.2}$$

such that the CLT (4.3) is not valid. This theorem shows that the functions $\cos 2\pi n_k \omega$ behave like independent random variables as long as $\{n_k\}$ satisfies (5.1) and at the critical speed (5.2) this behaviour turns into strong dependence. Several authors investigated the properties of $\cos 2\pi n_k \omega$ under

(5.1) or (4.6); on the other hand, very little is known in the strongly dependent domain. Berkes [3], [4], [5] investigated the asymptotic properties of $\cos 2\pi n_k\omega$ in the "critical domain", i.e. in the immediate neighbourhood of the speed (5.2). As it turned out, in this domain $\{\cos 2\pi n_k\omega\}$ has a number of interesting and unusual probabilistic properties, very different from classical i.i.d. behaviour: normed trigonometric sums $a_N^{-1}\sum_{k\leqq N}\cos 2\pi n_k\omega$ can converge to nongaussian limit distributions and the classical LIL (4.2) is replaced by a variety of "fractional" LIL properties, partly resembling the classical LIL, but containing various nonclassical "disturbing" terms. The simplest nongaussian limit theorem for $\cos 2\pi n_k\omega$ is the following result from Berkes [3]:

THEOREM G. *For any sufficiently slowly decreasing sequence $\varepsilon_k\downarrow 0$ there exists a sequence $\{n_k\}$ of positive integers such that*

$$n_{k+1}/n_k \geqq 1+\varepsilon_k/\sqrt{k} \qquad (k\geqq k_0)$$

and for some bounded centering sequence $\{d_N\}$ we have

$$\mathbf{P}\left\{\omega:\ \frac{1}{\sqrt{N\varepsilon_N}}\sum_{k\leqq N}\cos 2\pi n_k\omega - d_N < x\right\} \to \frac{1}{\pi}\int_{-\infty}^{x}\frac{1}{1+t^2}dt. \tag{5.3}$$

There are many other possible limit distributions of normed sums $a_N^{-1}\sum_{k\leqq N}\cos 2\pi n_k\omega$ in the critical domain: for example, in Berkes [3] a one-parameter class $\{F_A,\ 0<A<\infty\}$ of infinitely divisible limit distributions is constructed which connects, in a continuous way, the Gaussian and Cauchy distributions. Nongaussian behaviour of $\sum_{k\leqq N}\cos 2\pi n_k\omega$ is intimately connected with the irregularities of $n_k\omega$ mod 1 and the high discrepancies in (2.6) and (4.7). For example, the sequence $\{n_k\}$ for which (2.6) holds is identical with the sequence $\{n_k\}$ in Theorem G, satisfying the Cauchy limit distribution theorem (5.3). In the Hadamard lacunary case $\sum_{k\leqq t}\cos 2\pi n_k\omega$ behaves like a Wiener process (cf. Theorem B), i.e. its order of magnitude is $O(t\log\log t)^{1/2}$ and this leads to a classical LIL for the discrepancy $D_N(\omega)$ of $n_k\omega$. On the other hand, for the sequence $\{n_k\}$ in Theorem G the partial sum process $\sum_{k\leqq t}\cos 2\pi n_k\omega$ resembles a Cauchy process whose trajectories are discontinuous, fluctuating rather wildly. As a consequence, the process $\sum_{k\leqq t}\cos 2\pi n_k\omega$, while staying around $\sqrt{t}$ most of the time, occasionally starts climbing rapidly and reaches the order of magnitude $(t\log t)^{1/2}$ infinitely many times. In view of Koksma's inequality, this "Cauchy effect" leads to (2.6).

In analogy with the change of behaviour of $\cos 2\pi n_k\omega$ at the critical speed (5.2), it is natural to expect an analogous change in the LIL behaviour of $D_N(\omega)$ at (5.2). In this respect, we formulate the following

Conjecture. *Let $\{n_k\}$ be a sequence of integers satisfying the Erdős gap condition* (4.6). *Then*

$$D_N(\omega) = O(N^{-1/2}(\log\log N)^c) \qquad \text{a.s.}$$

for some constant $c > 0$.

In other words, we believe that Erdős' conjecture (2.4), while false for general $\{n_k\}$, is valid at least under (4.6). Recalling that there are sequences $\{n_k\}$ growing almost with the speed (5.2) such that (2.6) holds (cf. Remark 2 after Theorem A), the validity of the above conjecture would imply an essential change in the order of magnitude of $D_N(\omega)$ around the gap condition (5.2).

REFERENCES

[1] Baker, R. C., Metric number theory and the large sieve, *J. London Math. Soc.* (2) **24** (1981), 34–40. *MR* **83a**:10086

[2] Berkes, I., An almost sure invariance principle for lacunary trigonometric series, *Acta Math. Acad. Sci. Hungar.* **26** (1975), 209–220. *MR* **54** #14031

[3] Berkes, I., Non-Gaussian limit distributions of lacunary trigonometric series, *Canad. J. Math.* **43** (1991), 948–959. *MR* **92k**:60108

[4] Berkes, I., Critical LIL behavior of the trigonometric system, *Trans. Amer. Math. Soc.* **338** (1993), 553–585. *MR* **93j**:60035

[5] Berkes, I., An optimal condition for the LIL for trigonometric series, *Trans. Amer. Math. Soc.* **347** (1995), 515–530.

[6] Berkes, I. and Philipp, W., The size of trigonometric and Walsh series and uniform distribution mod 1, *J. London Math. Soc.* **50** (1995), 454–464.

[7] Carleson, L., On convergence and growth of partial sums of Fourier series, *Acta Math.* **116** (1966), 135–157. *MR* **33** #7774

[8] Cassels, J. W. S., Some metrical theorems of Diophantine approximation III, *Proc. Cambridge Philos. Soc.* **46** (1950), 219–225. *MR* **12**-162

[9] Dhompongsa, S., Uniform laws of the iterated logarithm for Lipschitz classes of functions, *Acta Sci. Math. (Szeged)* **50** (1986), 105–124. *MR* **87m**:26003

[10] Erdős, P., On trigonometric sums with gaps, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **7** (1962), 37–42. *MR* **26** #2797

[11] Erdős, P., Problems and results on diophantine approximations, *Compositio Math.* **16** (1964), 52–65. *MR* **31** #3382

[12] Erdős, P. and Gál, I. S., On the law of the iterated logarithm, I–II, *Nederl. Akad. Wetensch. Proc. Ser. A* **58** = *Indagationes Math.* **17** (1955), 65–76, 77–84. *MR* **16**-1016

[13] Erdős, P. and Koksma, J. F., On the uniform distribution modulo 1 of sequences $(f(n,\theta))$, *Nederl. Akad. Wetensch. Proc.* **52** (1949), 851–854 = *Indagationes Math.* **11** (1949), 299–302. *MR* **11**-331

[14] Hunt, R. A., On the convergence of Fourier series, *Orthogonal expansions and their continuous analogues* (Proc. Conf., Edwardsville, Ill., 1967), Southern Illinois Univ. Press, Carbondale, Ill., 1968, 235–255. *MR* **38** #6296

[15] Kesten, H., The discrepancy of random sequences $\{kx\}$, *Acta Arith.* **10** (1964/65), 183–213. *MR* **29** #5807

[16] Khintchine, A. J., Einige Sätze über Kettenbrüche, mit Anwendungen auf die Theorie der diophantischen Approximationen, *Math. Ann.* **92** (1924), 115–125. *Jb. Fortschritte Math.* **50**, 125

[17] KUIPERS, L. and NIEDERREITER, H., *Uniform distribution of sequences*, Pure and Applied Mathematics, Wiley, New York–London–Sydney, 1974. *MR* **54** #7415

[18] PHILIPP, W., Limit theorems for lacunary series and uniform distribution mod 1, *Acta Arith.* **26** (1974/75), 241–251. *MR* **52** #325

[19] PHILIPP, W., A functional law of the iterated logarithm for empirical distribution functions of weakly dependent random variables, *Ann. Probab.* **5** (1977), 319–350. *MR* **56** #1397

[20] PHILIPP, W., Empirical distribution functions and strong approximation theorems for dependent random variables. A problem of Baker in probabilistic number theory, *Trans. Amer. Math. Soc.* **345** (1994), 707–727.

[21] PHILIPP, W. and STOUT, W., Almost sure invariance principles for partial sums of weakly dependent random variables, *Mem. Amer. Math. Soc.* **2** (1975), issue 2, no. 161, 1975. *MR* **55** #6570

[22] RÉVÉSZ, P., *The laws of large numbers*, Probability and Mathematical Statistics, Vol. 4, Academic Press, New York–London, 1968. *MR* **39** #6391

[23] SALEM, R. and ZYGMUND, A., On lacunary trigonometric series, *Proc. Nat. Acad. Sci. U.S.A.* **33** (1947), 333–338. *MR* **9**–181

[24] SALEM, R. and ZYGMUND, A., La loi du logarithme itéré pour les séries trigonométriques lacunaires, *Bull. Sci. Math.* (2) **74** (1950), 209–224. *MR* **12**–605

[25] TAKAHASHI, S., On the law of the iterated logarithm for lacunary trigonometric series, *Tôhoku Math. J.* (2) **24** (1972), 319–329. *Zbl* **241**.42007

[26] TAKAHASHI, S., On the law of the iterated logarithm for lacunary trigonometric series II, *Tôhoku Math. J.* (2) **27** (1975), 391–403. *Zbl* **315**.40009

[27] TIJDEMAN, R., On integers with many small prime factors, *Compositio Math.* **26** (1973), 319–330. *MR* **48** #3896

[28] WEYL, H., Über die Gleichverteilung von Zahlen mod. Eins, *Math. Ann.* **77** (1916), 313–352. *Jb. Fortschritte Math.* **46**, 278–280

(Received February 23, 1995)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

DEPARTMENT OF STATISTICS
UNIVERSITY OF ILLINOIS
725 S. WRIGHT STREET
CHAMPAIGN, IL 61820
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 27–33

ON THE GEOMETRY OF CHARACTERISTIC CURVES

D. BSHOUTY and G. J. SZÉKELY*

Abstract

In this paper we study the geometric structure of ranges of characteristic functions of random variables.

1. Introduction

DEFINITION 1. A *characteristic curve* (c.c.) of a characteristic function (c.f.) f is the graph

$$K \equiv K_f = \{f(t) \in \mathbb{C} : t \in \mathbb{R}\}$$

where $\mathbb{R}$ denotes the set of real numbers and $\mathbb{C}$ is the complex plane.

A c.f. determines a unique c.c. but the converse is not true as the c.c. of all identically constant random variables is the same unit circle.

The basic properties of c.f.'s, namely, $f(0) = 1$, $|f(t)| \leqq 1$ and $f(-t) = \overline{f(t)}$ imply that $1 \in K_f$, K_f is a subset of the unit disk D and K_f is symmetric with respect to the real axis.

DEFINITION 2. A curve $K \subseteqq D = \{z \in \mathbb{C} : |z| \leqq 1\}$ will be called an *admissible curve* (a.c.) if $z = 1 \in K$ and K is symmetric with respect to the real axis.

The names admissible curve and characteristic curve are justified by the fact that $f(t)$ is a continuous function.

If a c.f. has the property $|f(t_0)| = 1$ for some $t_0 \neq 0$ then the corresponding distribution is a *lattice distribution* (l.d.) (see Lukacs [5] p. 18 Th. 2.1.4). C.f.'s of l.d.'s are closely related to analytic functions. Indeed, every such

1991 *Mathematics Subject Classification.* Primary 60E10; Secondary 30C99.

Key words and phrases. Characteristic function, Schwarz–Christoffel formula, geometric function theory, Riemann mapping theorem.

* This research started when both authors were visiting professors at Bowling Green State University, the second author the first Eugene Lukacs Professor. Since then the second author has been supported by the Hungarian National Foundation for Scientific Research Grant No. 1405, 1905 and T 016237.

c.f. has the representation

$$f(t) = \sum_{-\infty}^{\infty} a_n e^{i(n+\zeta_0)\alpha t}$$

for some ζ_0, $\alpha > 0$, and $a_n \geq 0$ for all $n \in \mathbb{Z}$. Hence

$$e^{-i\zeta_0 \alpha t} f(t) = \sum_{-\infty}^{\infty} a_n e^{in\alpha t} \equiv h(e^{i\alpha t})$$

where

$$h(z) = a_0 + \sum_{n=1}^{\infty} (a_n z^n + a_{-n}(\overline{z})^n) .$$

This $h(z)$ represents a harmonic function in the open unit disk U and $f(t)$ is related to the boundary values $h(\partial U)$. If the *lattice distribution* is *centered at zero* (l.d.z.), i.e. $\zeta_0 = 0$, $f(t)$ is the boundary of a harmonic function in U. If, furthermore, the support of the l.d.z. is in $N \cup \{0\}$, $f(t)$ is the boundary value of an analytic function in U.

Our "Geometry of Characteristic Curves" and the classical "Geometric Function Theory" (see [3]) are therefore two related topics with one major difference, namely, whereas in the first case we are interested in the boundary values, say, of an analytic function, in the second case we focus our attention on the image domain as a whole. A typical example of the second case is the celebrated Riemann Mapping Theorem which addresses the following question. Which domains in the plane are images of the unit disk under a one-to-one analytic function? Riemann's Theorem answers this question completely by characterizing these domains topologically. These images are exactly all simply connected domains in the plane whose boundaries contain more than one point. (The whole plane itself has to be excluded by Liouville's theorem claiming that every bounded entire function is constant.) In this context it is therefore appropriate to pose the following

PROBLEM 1. Characterize all admissible curves that are characteristic curves.

This problem turns out to be a very difficult one since we restrict the search among functions $h(z)$ with nonnegative coefficients a_n.

Another topic considered in Geometric Function Theory is labeled *"Covering Theorems"*. These are theorems for various classes of analytic functions that establish certain properties of sets that are entirely contained in the image domain of each function of the corresponding class. A typical theorem is the following:

For every function $h(z)$ analytic in U with $h(0) = h'(0) - I = 0$, $h(U)$ contains the disk $\left\{z : |z| < \frac{\sqrt{3}}{4}\right\}$.

In our case instead of intersection of all $h(U)$ where h is in a given class we can consider the union of all $f(U)$ where f is in a given class.

PROBLEM 2. Given a class S of c.f.'s (e.g. of unimodal, decreasing failure rate, etc.) describe the set $K_S = \bigcup_{f \in S} f(U)$.

In this paper Problem 1 will be discussed in detail. Concerning Problem 2 we only mention two simple results:

(i) the set of all Poisson c.c.'s cover the open unit disk except the point 0;

(ii) the set of c.c.'s of unimodal distributions with mode at 0 (they are exactly those distributions that have cumulative distribution functions convex on the negative half-line, and concave on the positive half-line) are covered by the disk $\{z \in \mathbb{C} : |z - 1/4| \leq 3/4\}$.

The first result is almost obvious and the second one can be deduced from Khintchin's well-known theorem claiming that every unimodal distribution with mode at 0 is a mixture (convex combination) of uniform distributions on finite intervals starting or ending at 0 ([4], [2] p. 158). Thus it suffices to show that the c.f.'s of all such uniform distributions satisfy the desired property, i.e. $\left|\frac{e^{itx}-1}{itx} - \frac{1}{4}\right| \leq \frac{3}{4}$ for all real nonzero x and t.

2. Hors-d'oeuvre

We shall see that the following statements hold:

(i) every regular n-gon with vertices $e^{2\pi k/n}$, $k = 0, 1, \ldots, n-1$ is a c.c. (of a l.d.z.);

(ii) if a triangle has an angle greater than $2\pi/3$ at $z = 1$, then it is not a c.c.

QUESTION 1. Does there exist a nonregular triangle which is a c.c. (preferably of a l.d.z.)?

It is clear that degenerate random variables (i.e. random variables taking one single value with probability 1) can be characterized by the property that their c.c.'s is the unit circle. (Suppose the random variable X has c.f. f, the random variable X' has distribution function $\mathbf{P}(X' < x) = 1 - \mathbf{P}(X < -x)$, i.e. the c.f. of X' is $\overline{f}$ thus $1 \equiv |f(t)| = f\overline{f}$ implies that if these random variables are independent then $X - X' \equiv 0$, i.e. $X = X'$ with probability 1, and this contradicts to independence except if $X = X'$ is degenerate.) Now, what if K is a circle but not the unit circle. All such admissible circles are c.c.'s of random variables taking only two values, and one of them is 0.

QUESTION 2. Is there a random variable supported on more than 2 values such that its c.c. is a circle?

Interestingly enough the c.c. of the exponential distribution whose c.f. is $(1 - it)^{-1}$ is almost a circle. Only the point $z = 0$ is missing. This shows that the c.c. of the random variable X with distribution $\mathbf{P}(X=0)=\mathbf{P}(X=1)=$ $=1/2$ is almost identical with the c.c. of the above mentioned exponential distribution (their closure is the same circle).

QUESTION 3. Is there a c.c. which covers the open unit disk U?

Thus we are looking for c.c.'s that are Peano curves. First we note that a simple shift of a random variable may completely destroy the shape of its c.c., it can easily become dense in the unit disk. For example the c.f. of the coin tossing random variable $\mathbf{P}(X=1)=\mathbf{P}(X=-1)=1/2$ is $f(t)=\cos t$ and its c.c. is the interval $[-1,1]$, while the c.f. of $\mathbf{P}(X=a\pm 1)=1/2$ is $f_a(t)=$ $=e^{iat}\cos t$ and the corresponding c.c. is everywhere dense in the unit disk if $2\pi/a$ is irrational $\bmod 2\pi$. To see this, put $\mathbb{R}_\vartheta=\{re^{i\vartheta}:0\leqq r\leqq 1\}$. Then it suffices to show that $\{f(t):t\in\mathbb{R}\}$ is dense in $\mathbb{R}_\vartheta$ for every ϑ. Indeed, for $t_k=(\vartheta+2\pi k)/a$ we have $f_a(t_k)=e^{i\vartheta}\cos[\vartheta+2\pi k]/a$ and here the cos factor is known to be dense in $[-1,1]$. Our argument implies a somewhat more general

PROPOSITION 1. *Let f be a c.f. of a lattice distribution supported on $\mathbb{Z}$. Then for $2\pi/a$ irrational* $\bmod 2\pi$ *the c.c. of the c.f. $e^{iat}f(t)$ is dense in the annulus* $\inf|f|\leqq|z|\leqq 1$.

A partial negative answer to our Question 3 follows from a problem by Pólya and Szegő ([8], p. 28, Problem 170). This problem claims that if f is an analytic function on the real axis, then the corresponding c.c. cannot be a Peano curve. More precisely the following proposition is true.

PROPOSITION 2. *There does not exist a function which is regular analytic along the real axis and which assumes for real values of the variable every value in the interior of a fixed circle. In short: There does not exist analytic Peano curve.*

3. Main results

In the following we shall focus our attention to nonintersecting closed curves. One might think that if K_f is a closed Jordan curve, then there exists a $t_0\neq 0$ such that $f(t_0)=1$, and thus f is the c.f. of a lattice distribution. This is not the case and examples can be constructed using a theorem by Shimitzu (see [6], p. 8). All these examples have the property that $0\notin \operatorname{Int} K_f$. We conjecture that if K_f is a closed Jordan curve and $0\in\operatorname{Int} K_f$ then f is the c.f. of a lattice distribution.

Another remark is the following. If a c.f. is smooth at the point $z=1$, then it is almost as smooth at every other point (see Lukacs [6] p. 22). This implies e.g. that not every convex a.c. is a c.c. On the other hand we shall see that every admissible regular polygon is a c.c. of a l.d.z.

As noted in the Introduction, if $K = h(\partial U)$ where h is an analytic function in the unit disk U with nonnegative Maclaurin coefficients, then K is the c.c. of a l.d.z. A particular case is

THEOREM 1. *Let K be a convex admissible curve, symmetric with respect to the imaginary axis. If the opening angle at $z = 1$ is less than $\pi/2$, then K is the characteristic curve of a lattice distribution supported on N.*

Theorem 1 is a special case of the more general

THEOREM 2. *Let K be a convex admissible curve, symmetric to the rotation of angle $2\pi/n$, $n \in N$, around the origin. If the opening angle at $z = 1$ is less than $(n-1)\pi/n$, then K is the characteristic curve of a lattice distribution supported on N.*

REMARK 1. The angle $(n-1)\pi/n$ in Theorem 2 is sharp. It cannot be replaced by any larger number. Boas [1] implies that if a c.c. of a l.d.z. is a polygon that has an opening angle $\beta < \pi$ at $z = 1$, then it has an opening angle at least β at every other point. This shows that no *triangle with opening angle $\beta > 2\pi/3$ at $z = 1$ can be a c.c.* We do not know if there exists a non-equilateral triangle that is a c.c. at all.

COROLLARY 1. *Every admissible regular n-gon that is symmetric with respect to rotations $2\pi/n$ around the origin is a characteristic curve of a lattice distribution supported on N.*

This Corollary and the following Proposition imply that the polygons of the Corollary are the only polygons with vertices on the unit circle that are c.c.'s.

PROPOSITION 3. *The points of a c.c. that are on the unit circle are equidistant.*

This Proposition is a simple consequence of Theorem 2.1.4 of Lukacs [5].

REMARK 2. The complexity of the characterization of c.c.'s (even of convex ones) is best seen by considering the class of convex hexagons that touch the unit circle only at $z = 1$. On one hand there exist non-regular convex hexagons that are c.c.'s, namely, consider a convex combination of the identically 1 c.f. and the c.f. whose c.c. is a hexagon that satisfies the premises of Theorem 2. On the other hand, using Proposition 3, we infer that there are hexagons that are not c.c.'s.

REMARK 3. If K is an admissible curve, star-like with respect to the origin, and symmetric with respect to the imaginary axis, then $K^2 = \{z^2 : z \in \in K\}$ is a c.c. The star-likeness of K implies the star-likeness of K^2, and the symmetry of K implies that K^2 is a Jordan curve.

The proof of Theorem 2 is similar to the proof of Theorem 1 and is omitted. We only have to prove Theorem 1.

4. Proof of Theorem 1

The main tool in this proof is the Schwarz–Christoffel formula (see e.g. Nehari [7]). First we assume that K is a convex admissible polygon symmetric with respect to the imaginary axis. Denote the exterior angles of the first quadrant by $\pi\alpha_1, \pi\alpha_2, \dots, \pi\alpha_{n-1}$, and the exterior angles at $z=1$ and at the imaginary axis at the point, say, ai by $\pi\alpha_0$ and $\pi\alpha_n$, respectively. By the symmetry properties of our domain Int K, the Schwarz–Christoffel mapping of the unit disk onto Int K given by Nehari p. 193 is

$$f(z)=c\int_0^z \frac{d\zeta}{(1-\zeta^2)^{\alpha_0}(1+a^2\zeta^2)^{\alpha_n}\prod_{i=1}^{n-1}(1-\eta_i^2\zeta^2)^{\alpha_i}(1-\overline{\eta_i^2}\zeta^2)^{\alpha_i}} \tag{1}$$

for some $c>0$, and η_i such that $|\eta_i|=1$ and $\alpha_i>0$ such that

$$4\sum_{i=1}^{n-1}\alpha_i+2(\alpha_0+\alpha_n)=2. \tag{2}$$

Here, η_i are the points of the unit circle that are mapped by f to the vertices of the polygon. In particular, $K=\{f(e^{it}):t\in\mathbb{R}\}$. We shall see that $f(e^{it})$ is a c.f. To see this we only need to show that the coefficients in the Maclaurin series of $f(z)$ are nonnegative. Since $f(0)=0$ we only have to show that the derivative $f'(z)$ has nonnegative coefficients except possibly for the constant coefficient. Since $f(0)=0$ and $f'(0)=c>0$, we only have to show that $f(z)/c$ has nonnegative coefficients. Now

$$\begin{aligned}
&\log\{f'(z)/c\}=\\
&\quad=-\alpha_0\log(1-z^2)-\alpha_n\log(1+a^2z^2)-\\
&\qquad-\sum_{i=1}^{n-1}\alpha_i\left[\log(1-\eta_i^2z^2)+\log(1-\overline{\eta_i^2}z^2)\right]=\\
&\quad=\sum_{k=1}^{\infty}\left(\alpha_0+2\sum_{i=1}^{n-1}\alpha_i\mathrm{Re}\,\eta_i^k+\alpha_n(-a^2)^k\right)\frac{z^{2k}}{k}.
\end{aligned}$$

By assumption $\alpha_0\geqq 1/2$, and hence by (2) $\alpha_n+2\sum_{i=1}^{n-1}\alpha_i\leqq 1/2$ thus

$$\alpha_0+2\sum_{i=1}^{n-1}\alpha_i\mathrm{Re}\,\eta_i^k+\alpha_n(-a^2)^k\geqq\alpha_0-2\sum_{i=1}^{n-1}\alpha_i-\alpha_n\geqq 0.$$

The result follows for this case (for polygons).

If K is not a polygon, then it can be approximated by an increasing sequence of polygonal c.c.'s K_n. For each K_n there corresponds a c.f. f_n of the form (1). The family $\{f_n(z)\}_{n=1}^{\infty}$ is a family of Riemann Mappings of $\operatorname{Int} K_n$ that converges in the kernel sense to $\operatorname{Int} K$ (see [3], p. 54), hence f_n converges locally uniformly to f that maps the unit disk onto $\operatorname{Int} K$. Moreover, ([3], p. 41) $K = \{f(e^{it}) : t \in \mathbb{R}\}$. Since the coefficients of f_n were nonnegative, and each coefficient of f is the limit of coefficients of f_n the coefficients of f are nonnegative. The proof of Theorem 1 is complete.

Finally, the authors thank Professor Sz. Gy. Révész for many valuable comments and suggestions.

REFERENCES

[1] Boas, R. P., Lipschitz behaviour and integrability of characteristic functions, *Ann. Math. Statist.* **38** (1967), 32–36. *MR* **34** # 6822

[2] Feller, W., *An introduction to probability theory and its applications*, Vol. II, 2nd edition, Wiley, New York, 1971. *MR* **42** # 5292

[3] Goluzin, G. M., *Geometric theory of functions of a complex variable*, Translations of Mathematical Monographs, Vol. 26, American Mathematical Society, Providence, RI, 1969. *MR* **40** # 308

[4] Khintchin, A. Ya., On unimodal distributions, *Izv. Naučno-Issled. Inst. Mat. i Meh. Tomsk. Gos. Univ. im V. V. Kujbyševa* **2** (1938), no. 2, 1–7 (in Russian). *Jb. Fortschritte Math.* **64**, 1197

[5] Lukacs, E., *Characteristic functions*, 2nd edition, revised and enlarged, Hafner Publ. Co., New York, 1970. *MR* **49** # 11595

[6] Lukacs, E., *Developments in characteristic function theory*, Macmillan, New York, 1983. *MR* **86m**:60047

[7] Nehari, Z., *Conformal mapping*, McGraw–Hill, New York–Toronto–London, 1952. *MR* **13**–640

[8] Pólya, G. and Szegő, G., *Problems and theorems in analysis*, Vol. II, Theory of functions, zeros, polynomials, determinants, number theory, geometry, Springer Study Edition, Springer-Verlag, New York–Heidelberg, 1976. *MR* **57** # 5529

(Received September 26, 1994)

DEPARTMENT OF MATHEMATICS
TECHNION
ISRAEL INSTITUTE OF TECHNOLOGY
IL–32000 HAIFA
ISRAEL

BUDAPESTI MŰSZAKI EGYETEM
ÉPÍTŐMÉRNÖKI KAR
MATEMATIKA TANSZÉK
H ÉPÜLET, V. EM. 7
STOCZEK U. 2
H–1521 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 35–46

ON THE LENGTH OF THE LONGEST MONOTONE BLOCK

E. CSÁKI[1] and A. FÖLDES[2]

Dedicated to Pál Révész on the occasion of his 60th birthday, who encouraged and helped us in this work

Abstract

The length of the longest monotone block is studied. It is shown that this length is of order $\log n$ for any discrete distribution. On the other hand, the length of the longest strictly monotone block depends on the distribution. As examples, we discuss the case of geometric and Poisson distribution.

1. Introduction

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with common distribution function $F(x)$. If

$$X_{i+1} \leqq X_{i+2} \leqq \ldots \leqq X_{i+j} \tag{1.1}$$

then $X_{i+1}, \ldots, X_{i+j}$ form a monotone block (MB) of length j. If

$$X_{i+1} < X_{i+2} < \ldots < X_{i+j} \tag{1.2}$$

then $X_{i+1}, \ldots, X_{i+j}$ form a strictly monotone block (SMB) of length j. It is interesting to investigate the length $\mu(n)$ of the longest MB or the length of the longest SMB $\mu^S(n)$ in the first n trials. In case of a continuous distribution function $F(x)$, this problem was raised by Pittel [11]. He proved that

$$\lim_{n\to\infty} \frac{\log\log n}{\log n}\mu(n) = 1 \quad \text{a.s.}. \tag{1.3}$$

Obviously, for continuous $F(x)$ to consider the longest MB or SMB is the same problem, that is to say, for a continuous $F(x)$

$$\mathbf{P}(\mu(n) = \mu^S(n)) = 1.$$

Révész [12] gave a more precise result than (1.3):

1991 *Mathematics Subject Classification*. Primary 60F15; Secondary 60F05.

Key words and phrases. Monotone block, discrete distribution.

[1] Research supported by Hungarian National Foundation for Scientific Research Grant No. 1905.

[2] Research supported by a PSC CUNY Grant No 6-65342.

THEOREM (Révész [12]). *Let b_n be the solution of the equation*

$$b_n e^{b_n} = (\log n)e^{-1},$$

$$f(n) = \frac{\log n}{b_n} - \frac{1}{2}, \quad \alpha(n) = f(n) - [f(n)]$$

$$l_n(\epsilon) = \begin{cases} [f(n)] - 3, & \text{if } \alpha(n) \leqq \epsilon \\ [f(n)] - 2, & \text{if } \alpha(n) > \epsilon \end{cases}$$

$$u_n(\epsilon) = \begin{cases} [f(n)] + 2, & \text{if } \alpha(n) \leqq 1 - \epsilon \\ [f(n)] + 3, & \text{if } \alpha(n) > 1 - \epsilon \end{cases}$$

$$v_n(\epsilon) = \begin{cases} [f(n)], & \text{if } \alpha(n) \leqq \epsilon \\ [f(n)] + 1, & \text{if } \alpha(n) > \epsilon \end{cases}$$

$$j_n(\epsilon) = \begin{cases} [f(n)] - 1, & \text{if } \alpha(n) \leqq 1 - \epsilon \\ [f(n)], & \text{if } \alpha(n) > 1 - \epsilon. \end{cases}$$

Then for $\mu(n, C)$ the length of the longest MB for an arbitrary continuous distribution, and for any $\epsilon > 0$

$$u_n(\epsilon) \in UUC(\mu(n, C))$$
$$v_n(\epsilon) \in ULC(\mu(n, C))$$
$$j_n(\epsilon) \in LUC(\mu(n, C))$$
$$l_n(\epsilon) \in LLC(\mu(n, C)).$$

For the definitions of the classes UUC, ULC, LUC, LLC we refer to Révész [13]. Further results were given by Grill [6] and Novak [9].

However, similar problems for the case of noncontinuous $F(x)$ remained open. In this paper we will investigate the case of discrete $F(x)$. The length of the longest MB and SMB with unspecified discrete distribution will be denoted by $\mu(n, F)$ and $\mu^S(n, F)$, respectively. For the geometric distribution the exact probability of $\mathbf{P}(X_1 \leqq X_2 \leqq \ldots \leqq X_k)$ and $\mathbf{P}(X_1 < X_2 < \ldots < X_k)$ will be calculated. As a consequence, the limit distribution of the waiting time $\tau(m, G)$ for an MB of length m, and the waiting time $\tau^S(m, G)$ for an SMB of length m will be given. The corresponding strong law for $\mu(n, G)$ and $\mu^S(n, G)$ will be formulated. (Here G stands for the geometric distribution.) Moreover we will show, that for any discrete distribution $F(x)$, $\mu(n, F)$ behaves as $\mu(n, G)$ with an appropriately selected success probability p. However, we will show, that the length of the longest SMB varies with the distribution. As an example the strong law for $\mu(n, \mathbf{P})$, the length of the longest SMB for the Poisson distribution will be presented.

The problem of the longest (strictly) monotone block is strongly related to the problem of the longest head run. The length $m(n)$ of the longest head run of n Bernoulli trials $\mathbf{P}(X_i = 1) = \mathbf{P}(X_i = -1) = \frac{1}{2}$ was raised by T. Varga, first answered by Erdős–Rényi [3], and a surprisingly precise answer was given by Erdős–Révész [4].

THEOREM (Erdős–Révész [4]). *Let $\{a_n\}$ be an increasing sequence of positive numbers, and let*

$$A(\{a_n\}) = \sum_{n=1}^{\infty} 2^{a_n}.$$

Then

$$\begin{aligned} a_n &\in UUC(m(n)) \quad \text{if} \quad A(\{a_n\}) < \infty \\ a_n &\in ULC(m(n)) \quad \text{if} \quad A(\{a_n\}) = \infty \end{aligned}$$

and for any $\epsilon > 0$

$$\begin{aligned} [\lg n - \lg\lg\lg n + \lg\lg e - 1 + \epsilon] &\in LUC(m(n)) \\ [\lg n - \lg\lg\lg n + \lg\lg e - 2 + \epsilon] &\in LLC(m(n)) \end{aligned}$$

*(*lg *stands for logarithm with base* 2*).*

If the Bernoulli variable is taken to be $\mathbf{P}(X_i = 1) = p$, $\mathbf{P}(X_i = -1) = 1 - p$, then similar results are valid. We only mention that in this case the leading term changes for $\log_{\frac{1}{p}} n$. ($\log_a$ denotes logarithm with base a.)

Further results on this topic were obtained by Guibas and Odlyzko [7], Samarova [14], Deheuvels [2]. A very general method was worked out by Csáki–Földes–Komlós [1], Novak [9] and Móri [8] which can be applied in numerous situations including the longest head run, and monotone block problems, and will be applied in this work as well.

2. Results for the geometric distribution

LEMMA. *Let $\{X_i\}_{i=1}^{\infty}$ be an i.i.d. sequence of geometric random variables with*

$$\mathbf{P}(X_1 = j) = q^{j-1}p, \qquad j = 1, 2, \ldots \tag{2.1}$$

where $0 < p = 1 - q < 1$. Then

$$\mathbf{P}(X_1 \leqq X_2 \leqq \ldots \leqq X_k) = \frac{p^k}{\prod_{l=1}^{k}(1 - q^l)}, \quad k = 1, 2, \ldots \tag{2.2}$$

and

$$\mathbf{P}(X_1 < X_2 < \ldots < X_k) = \frac{p^k}{\prod_{l=1}^{k}(1 - q^l)} q^{\frac{k(k-1)}{2}}, \quad k = 1, 2, \ldots . \tag{2.3}$$

PROOF. Put

(2.4)
$$f_k = \mathbf{P}(X_1 \leqq X_2 \leqq \ldots \leqq X_k) = \sum_{j=1}^{\infty} \mathbf{P}(X_1 = j)\mathbf{P}(j \leqq X_2 \leqq X_3 \leqq \ldots \leqq X_k).$$

Then

(2.5)
$$\mathbf{P}(j \leqq X_2 \leqq X_3 \leqq \ldots \leqq X_k) =$$
$$= \mathbf{P}(X_2 - j + 1 \leqq \cdots \leqq X_k - j + 1 \mid X_2 \geqq j, \ldots, X_k \geqq j)\,(\mathbf{P}(X_2 \geqq j))^{k-1} =$$
$$= \mathbf{P}(X_2 \leqq \cdots \leqq X_k) q^{(j-1)(k-1)} = q^{(j-1)(k-1)} f_{k-1},$$

since the conditional distribution of $X_i - j + 1$ under the condition $X_i \geqq j$ is also geometric. Hence

(2.6)
$$f_k = \sum_{j=1}^{\infty} q^{j-1} p q^{(j-1)(k-1)} f_{k-1} = p f_{k-1} \sum_{j=1}^{\infty} q^{k(j-1)} = \frac{p}{1-q^k} f_{k-1}.$$

Being $f_1 = 1$, we get

(2.7)
$$f_k = \prod_{j=2}^{k} \left(\frac{p}{1-q^j}\right) = \frac{p^k}{\prod_{j=1}^{k}(1-q^j)}$$

proving (2.2). To get (2.3) we only have to repeat the previous argument establishing

(2.8)
$$\mathbf{P}(j < X_2 < X_3 < \cdots < X_k) = q^{j(k-1)} \mathbf{P}(X_2 < X_3 < \cdots < X_k)$$

instead of (2.5), and finish the proof as above.

To exploit the above result, we will apply our main lemma in Csáki–Földes–Komlós [1]. To formulate this result we need some notations. We only give the stationary form of the lemma, as this is the form we need here. Let $X_1, X_2, \ldots$ be a sequence of independent random variables, and let $\mathcal{F}_{n,m}$ denote the σ-algebra generated by the blocks of variables $X_n, X_{n+1}, \ldots, X_{n+m-1}$. $A_{n,m}$ will denote a sequence of events, where $A_{n,m} \in \mathcal{F}_{n,m}$. In the notation, m will be suppressed $A_{n,m}$ will be simply denoted by A_n. The purpose of the lemma is to find good approximations to the probabilities

(2.9)
$$\mathbf{P}\left(\bigcup_{1}^{n} A_i\right) = 1 - \mathbf{P}(\bar{A}_1 \ldots \bar{A}_n)$$

or find the limit distribution of the random variable

(2.10)
$$\tau(m) = \{\text{first } n \text{ such that } A_n \text{ occurs}\}.$$

MAIN LEMMA ([1]). *Assume that A_n is stationary (m is fixed), and there is a number α, $0 < \alpha \leqq 1$, such that the following three conditions hold for some $k \leqq m$, $\epsilon > 0$*

$$\tag{2.11} \begin{gathered} |\,\mathbf{P}(\bar{A}_2 \dots \bar{A}_k \mid A_1) - \alpha \,| < \epsilon, \\ \sum_{k \leqq i \leqq 2m} \mathbf{P}(A_i \mid A_1) < \epsilon, \\ \mathbf{P}(A_1) < \frac{\epsilon}{m}. \end{gathered}$$

Then for all $N > 1$

$$e^{-(\alpha+10\epsilon)N\mathbf{P}(A_1) - 2m\mathbf{P}(A_1)} < \mathbf{P}(\bar{A}_1 \dots \bar{A}_N) < e^{-(\alpha-10\epsilon)N\mathbf{P}(A_1) + 2m\mathbf{P}(A_1)}.$$

Consequently if

$$\lim_{m\to\infty} n(m)\mathbf{P}(A_1) = \lambda,$$

then

$$\tag{2.12} \lim_{m\to\infty} \mathbf{P}(\bar{A}_1\bar{A}_2 \dots \bar{A}_{n(m)}) = e^{-\alpha\lambda}.$$

We will apply this lemma for $\{X_i\}_{i=1}^{\infty}$ i.i.d.r.v.-s, so A_n will be stationary. Define the events A_n and A_n^S as follows:

$$\tag{2.13} A_n = \{X_n \leqq X_{n+1} \leqq \dots \leqq X_{n+m-1}\},$$

$$\tag{2.14} A_n^S = \{X_n < X_{n+1} < \dots < X_{n+m-1}\},$$

that is to say, if A_n or A_n^S occur, then the block of length m starting with X_n is MB or SMB, respectively. Denote by $\tau(m,G)$ and $\tau^S(m,G)$ the waiting time (see (2.10)) for an MB and for an SMB of length m, respectively. By elementary calculations, one can check, that all conditions of the main lemma are met for A_n and for A_n^S as well. (For A_n, $\alpha = q$ and for A_n^S, $\alpha = 1$.) Application of the main lemma combined with Lemma 2.1 leads to

THEOREM 2.1. *If $\{X_i\}_{i=1}^{\infty}$ is i.i.d. geometric with $\mathbf{P}(X_1 = k) = pq^{k-1}$, $k = 1, 2, \dots$, then for all $x \geqq 0$*

$$\lim_{m\to\infty} \mathbf{P}\left(\tau(m,G) > \frac{xC(m)}{qp^m}\right) = e^{-x}$$

where

$$C(m) = \prod_{k=1}^{m} (1 - q^k).$$

Theorem 2.1 easily implies the following strong law for $\mu(n, G)$.

THEOREM 2.2. *If* $\{X_i\}_{i=1}^{\infty}$ *is i.i.d. geometric with* $\mathbf{P}(X_1 = k) = pq^{k-1}$, $k = 1, 2, \ldots$, *then*

$$\lim_{n\to\infty} \frac{\mu(n,G)}{\log_{\frac{1}{p}} n} = 1 \quad a.s..$$

Similarly for the SMB case, we have

THEOREM 2.3. *If* $\{X_i\}_{i=1}^{\infty}$ *is i.i.d. geometric with* $\mathbf{P}(X_1 = k) = pq^{k-1}$, $k = 1, 2, \ldots$, *then for all* $x \geqq 0$

$$\lim_{m\to\infty} \mathbf{P}\left(\tau^S(m,G) > \frac{xC(m)}{p^m q^{\frac{m(m-1)}{2}}}\right) = e^{-x}.$$

THEOREM 2.4. *If* $\{X_i\}_{i=1}^{\infty}$ *is i.i.d. geometric with* $\mathbf{P}(X_1 = k) = pq^{k-1}$, $k = 1, 2, \ldots$, *then*

$$\lim_{n\to\infty} \frac{\mu^S(n,G)}{\sqrt{2\log_{\frac{1}{q}} n}} = 1 \quad a.s..$$

REMARK. Theorems 2.1 and 2.3 combined with the usual Borel–Cantelli arguments would provide much more precise upper and lower class results, which we omit.

Comparing Theorems 2.2. and 2.4 shows that the length of the longest SMB is much shorter than the longest MB, which is not surprising at all.

$\mu(n,G)$ and $\mu^S(n,G)$ do not have limiting distributions. Applying, however, Theorem 3.1 of Móri [8] one can easily prove the following results.

THEOREM 2.5. *If* $\{X_i\}_{i=1}^{\infty}$ *is i.i.d. geometric with* $\mathbf{P}(X_1 = k) = pq^{k-1}$, $k = 1, 2, \ldots$, *then for all real* t

$$\frac{1}{\log n}\sum_{i=1}^{n} \frac{1}{i} I(\mu(i,G) - \log_{\frac{1}{p}} i < t) = \int_t^{t+1} \exp\{-\frac{q}{c}p^z\}\, dz,$$

where $I(\cdot)$ *denotes the indicator and* $c = C(\infty) = \lim_{m\to\infty} C(m)$.

THEOREM 2.6. *If* $\{X_i\}_{i=1}^{\infty}$ *is i.i.d. geometric with* $\mathbf{P}(X_1 = k) = pq^{k-1}$, $k = 1, 2, \ldots$, *then for all real* t

$$\frac{1}{\log n}\sum_{i=1}^{n} \frac{1}{i} I(\mu^S(i,G) - g(i) < t) = \begin{cases} 0 & \text{if } t \leqq -1, \\ 1+t & \text{if } -1 < t < 0, \\ 1 & \text{if } 0 \leqq t, \end{cases}$$

where

$$g(i) = -\log_{\frac{1}{q}} \frac{1}{p} + \frac{1}{2} + \sqrt{\left(\log_{\frac{1}{q}} \frac{1}{p} - \frac{1}{2}\right)^2 + 2\log_{\frac{1}{q}} \frac{i}{c}}\,.$$

From Theorem 2.2 in Novak [9] we can obtain a rate of convergence result:

THEOREM 2.7. *If* $\{X_i\}_{i=1}^{\infty}$ *is i.i.d. geometric with* $\mathbf{P}(X_1 = k) = pq^{k-1}$, $k = 1, 2, \ldots$, *then*

$$\max_{1 \leqq k \leqq n} \left| \mathbf{P}(\mu(n, G) < k) - \exp\left\{ -\frac{nqp^{k+1}(1 - q^{k+1})}{C(k+2)} \right\} \right| = O\left(\frac{\log n}{n}\right)$$

as $n \to \infty$.

3. Monotone blocks with arbitrary discrete distribution

To investigate the case of an arbitrary discrete distribution, we might suppose without loss of generality, that the random variable takes the values $1, 2, 3, \ldots$ with probabilities $p_1, p_2, p_3, \ldots$. Introduce the notation

$$p = \max_{1 \leqq i < +\infty} p_i. \tag{3.1}$$

LEMMA 3.1. *If the discrete distribution has finite support with*

$$K = \max\{i : p_i > 0\} \tag{3.2}$$

we have for all $l = 1, 2, \ldots$

$$p^l \leqq \mathbf{P}(X_1 \leqq X_2 \leqq \ldots \leqq X_l) \leqq p^l \binom{l + K - 1}{K - 1}. \tag{3.3}$$

PROOF. The lower bound is obvious. The upper bound follows easily from

$$\mathbf{P}(X_1 \leqq X_2 \leqq \ldots \leqq X_l) =$$
$$= \sum_{1 \leqq i_1 \leqq \cdots \leqq i_l \leqq K} p_{i_1} p_{i_2} \cdots p_{i_l} \leqq p^l \binom{l + K - 1}{K - 1}.$$

LEMMA 3.2. *If the discrete distribution has infinite support, and*

$$K = \min\left\{ k : \sum_{j=k}^{\infty} p_j < p \right\} \tag{3.4}$$

we have for all $l = 1, 2, \ldots$

$$p^l \leqq \mathbf{P}(X_1 \leqq X_2 \leqq \cdots \leqq X_l) \leqq p^l \binom{l + K}{K}. \tag{3.5}$$

PROOF. Being the lower bound obvious, we prove only the upper bound. Using simple combinatorial argument, we get

$$
\begin{aligned}
&\mathbf{P}(X_1 \leqq X_2 \leqq \ldots \leqq X_l) = \mathbf{P}(K \leqq X_1 \leqq X_2 \leqq \ldots \leqq X_l) + \\
&+ \sum_{j=1}^{l} \mathbf{P}(X_1 \leqq X_2 \leqq \ldots \leqq X_j \leqq K \leqq X_{j+1} \leqq \ldots \leqq X_l) \leqq \\
&\leqq p^l + \sum_{j=1}^{l} \binom{j+K-1}{K-1} p^j p^{l-j} = p^l \sum_{j=0}^{l} \binom{j+K-1}{K-1} = p^l \binom{l+K}{K}.
\end{aligned} \tag{3.6}
$$

Combining Lemma 3.1 and 3.2 we might conclude that for every discrete distribution there exists a (big enough) K (depending on the distribution) such that for l big enough

$$
p^l \leqq \mathbf{P}(X_1 \leqq X_2 \leqq \ldots \leqq X_l) \leqq l^K p^l. \tag{3.7}
$$

(3.7) leads to the following strong law:

THEOREM 3.1. *Let $X_1, X_2, \ldots X_n \ldots$ be a sequence of i.i.d.r.v.-s with arbitrary discrete distribution. Then for the length of the longest MB $\mu(n, D)$ (D for discrete)*

$$
\lim_{n \to \infty} \frac{\mu(n, D)}{\log_{\frac{1}{p}} n} = 1 \quad a.s. \,. \tag{3.8}
$$

PROOF. To prove the upper part one has to show that for a convenient subsequence $\{n_k\}$

$$
\sum_{k=1}^{\infty} \mathbf{P}(\mu(n_k, D) > (\log_{\frac{1}{p}} n_{k-1})(1+\epsilon)) < \infty. \tag{3.9}
$$

Take $n_k = k^\alpha$ with a big enough α, such that $\alpha\epsilon > 1$, and apply the trivial upper bound

$$
\mathbf{P}(\mu(n, D) > l) \leqq n\mathbf{P}(X_1 \leqq X_2 \leqq \ldots \leqq X_l).
$$

Elementary calculations and (3.7) give (3.9).

To prove the lower part we have to show that

$$
\sum_{k=1}^{\infty} \mathbf{P}(\mu(n, D) < (\log_{\frac{1}{p}} n)(1-\epsilon)) < \infty. \tag{3.10}
$$

Denote

$$
A(i, m) = \{X_{im} \leqq X_{im+1} \leqq \cdots \leqq X_{(i+1)m-1}\},
$$

then clearly

$$\{\mu(n,D)<m\}\subset\bigcap_{i=1}^{[\frac{n}{m}]}A(i,m), \tag{3.11}$$

hence

$$\mathbf{P}(\mu(n,D)<m)\leqq(1-\mathbf{P}(A(1,m)))^{[\frac{n}{m}]}. \tag{3.12}$$

(3.12) combined with elementary calculations leads to (3.10). Since the above argument is well-known we omit the details.

Comparing Theorems 2.2 and 3.1 we arrive to the following conclusion. Let D be an arbitrary discrete distribution with maximal probability p defined by (3.1). Furthermore let G be geometric with success probability p. Then

$$\lim_{n\to\infty}\frac{\mu(n,D)}{\mu(n,G)}=1 \quad \text{a.s.}.$$

Finally observe that our result also implies that if $m(n)$ is the length of the longest head run in n independent Bernoulli trials, where $\mathbf{P}(X_i=1)=p$, then we also have

$$\lim_{n\to\infty}\frac{\mu(n,D)}{m(n)}=1 \quad \text{a.s.}.$$

4. Strictly monotone blocks for the Poisson distribution

To illustrate that the length of the longest SMB varies with the distribution, we briefly discuss the case of the Poisson distribution.

REMARK 4.1. If the discrete distribution has finite support with K defined by (3.2), then of course $\mu^S(n)=K$ for large n, but here we want to see a nontrivial example which differs from the geometric case.

LEMMA 4.1. *Let $\{X_i\}_{i=1}^{\infty}$ be an i.i.d. sequence of random variables with*

$$p_k=\mathbf{P}(X_1=k)=\frac{\lambda^k}{k!}e^{-\lambda}, \qquad k=0,1,2\dots.$$

Then

$$\mathbf{P}(X_1<X_2<\cdots<X_l)=e^{-\frac{l^2}{2}\log l(1+o(1))} \tag{4.1}$$

as $l\to\infty$.

PROOF. To get the lower bound, observe that

$$\mathbf{P}_l=\mathbf{P}(X_1<X_2<\cdots<X_l)>p_0p_1\dots p_{l-1}.$$

Using Stirling formula, elementary computations lead to
(4.2)

$$\mathbf{P}_l > \exp\left\{ -\sum_{j=1}^{l-1} j\log j - \lambda l + \frac{l(l-1)}{2}\log\lambda + \sum_{1}^{l-1} j - \frac{1}{2}\sum_{j=1}^{l-1}\log j - Cl \right\}$$

(where C is a conveniently selected constant). As

$$\sum_{j=1}^{l-1} j\log j < (\log l)\sum_{j=1}^{l-1} j = (\log l)\frac{l(l-1)}{2} \tag{4.3}$$

and all the other terms in the exponent of (4.2) are of smaller order, we have the lower bound.

For the upper bound observe that

$$\mathbf{P}_l \leqq \mathbf{P}(X_2 \geqq 1)\mathbf{P}(X_3 \geqq 2)\dots\mathbf{P}(X_l \geqq l-1). \tag{4.4}$$

Using the obvious inequality

$$\mathbf{P}(X_1 \geqq j) \leqq \left(\frac{e\lambda}{j}\right)^j \qquad j = 1, 2, \dots \tag{4.5}$$

(easily proved by Stirling formula) we get

$$\mathbf{P}_l \leqq \prod_{j=1}^{l-1}\left(\frac{e\lambda}{j}\right)^j \leqq \exp\left\{ -\sum_{j=1}^{l-1} j\log j + \frac{l(l-1)}{2}\log\lambda + \frac{l(l-1)}{2} \right\}. \tag{4.6}$$

Now as

$$\sum_{k=2}^{l-1} k\log k \geqq \int_1^l x\log x\,dx = \frac{l^2\log l}{2}(1+o(1))$$

and the other terms in the exponent of (4.6) are of smaller order, we get (4.1).

THEOREM 4.1. *For the length $\mu^S(n,\mathbf{P})$ of the longest SMB of a Poisson distribution*

$$\lim_{n\to\infty} \frac{\mu^S(n,\mathbf{P})}{2\sqrt{\frac{\log n}{\log\log n}}} = 1 \quad a.s. .$$

The proof of Theorem 4.1 runs like that of Theorem 3.1, except that we use Lemma 4.1 instead of (3.7), hence we omit it.

REMARK 4.2. Theorem 4.1 shows that the longest SMB in the Poisson case is shorter than in the geometric case, and that the leading term of the asymptotic of the length does not depend on λ. (4.2) and (4.6), however, together imply that

$$\mathbf{P}(X_1 < X_2 < \cdots < X_l) =$$
$$= \exp\left\{-\frac{l^2 \log l}{2} + l^2\left(\frac{3}{4} + \frac{1}{2}\log\lambda\right) + O(l\log l)\right\},$$

from which one can obtain the more precise result

$$\mu^S(n, \mathbf{P}) = 2\sqrt{\frac{\log n}{\log\log n}}\left(1 + \frac{\log\log\log n}{2\log\log n} + \frac{3/2 + \log\lambda - \log 2 + o(1)}{\log\log n}\right) \quad \text{a.s.}$$

as $n \to \infty$, showing that λ appears in the third term only.

Remark 4.3. Comparing our result with the continuous case, it turns out, that for any discrete distribution, the longest MB is longer than the longest MB for any continuous distribution. However, in our examples treated above, the longest SMB sequence was shorter than $\frac{\log n}{\log\log n}$ which is the length of the longest (S)MB for continuous distribution. T. Móri kindly provided us with the following example, where the longest SMB is of the same order as for continuous distribution.

$$\mathbf{P}_l = \mathbf{P}(X_1 < X_2 < \cdots < X_l) = \frac{1}{l!}\mathbf{P}(X_1, X_2, \ldots, X_l \text{ are different}),$$

hence, assuming that $p_1 \geqq p_2 \geqq \ldots$, we get

$$\frac{1}{l!}s_1 s_2 \ldots s_{l-1} \leqq \mathbf{P}_l \leqq \frac{1}{l!},$$

where $s_i = \sum_{j=i+1}^{\infty} p_j$. If

$$p_k = \frac{C}{(k+1)\log^2(k+1)}, \qquad k = 1, 2, \ldots$$

then

$$s_1 s_2 \ldots s_{l-1} = \exp\{O(l\log\log l)\},$$

hence

$$\mathbf{P}_l = \exp\{-l\log l(1 + o(1))\}$$

as $l \to \infty$ giving the same order of magnitude for SMB as for continuous distribution.

Acknowledgement. The authors are indebted to T. Móri for his fruitful remarks.

REFERENCES

[1] Csáki, E., Földes, A. and Komlós, J., Limit theorems for Erdős–Rényi type problems, *Studia Sci. Math. Hungar.* **22** (1987), 321–332. *MR* **89b**:60052

[2] Deheuvels, P., On the Erdős–Rényi theorem for random fields and sequences and its relationship with the theory of runs and spacings, *Z. Wahrsch. Verw. Gebiete* **70** (1985), 91–115. *MR* **86k**:60045

[3] Erdős, P. and Rényi, A., On a new law of large numbers, *J. Analyse Math.* **23** (1970), 103–111. *MR* **42** #6907

[4] Erdős, P. and Révész, P., On the length of the longest head-run, *Topics in information theory* (Second Colloq., Keszthely, 1975), Colloq. Math. Soc. János Bolyai, Vol. **16**, North-Holland, Amsterdam, 1977, 219–228. *MR* **57** #17788

[5] Földes, A., The limit distribution of the length of the longest head-run, *Period. Math. Hungar.* **10** (1979), 301–310. *MR* **81b**:60022b

[6] Grill, K., Erdős–Révész type bounds for the length of the longest run from a stationary mixing sequence, *Probab. Theory Related Fields* **75** (1987), 77–85. *MR* **88f**:60057

[7] Guibas, L. J. and Odlyzko, A. M., Long repetitive patterns in random sequences, *Z. Wahrsch. Verw. Gebiete* **53** (1980), 241–262. *MR* **81m**:60047

[8] Móri, T. F., The a.s. limit distribution of the longest head-run, *Canad. J. Math.* **45** (1993), 1245–1262.

[9] Novak, S. Yu., Longest runs in a sequence of m-dependent random variables, *Probab. Theory Related Fields* **91** (1992), 269–281. *MR* **93b**:60046

[10] Pittel, B. G., A process of runs and its convergence to the Brownian motion, *Stochastic Process. Appl.* **10** (1980), 33–48. *MR* **82a**:60044

[11] Pittel, B. G., Limiting behaviour of a process of runs, *Ann. Probab.* **9** (1981), 119–129. *MR* **82f**:60058

[12] Révész, P., Three problems on the length of increasing runs, *Stochastic Process. Appl.* **15** (1983), 169–179. *MR* **84h**:60064

[13] Révész, P., *Random walk in random and non-random environments*, World Scientific, Singapore, 1990. *MR* **92c**:60096

[14] Samarova, S. S., On the length of the longest head-run for a Markov chain with two states, *Teor Veroyatnost. i Primenen.* **26** (1981), 510–520 (in Russian). *MR* **83a**:60104. See also *Theor. Probab. Appl.* **26** (1981), 499–509.

(Received September 7, 1994)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

COLLEGE OF STATEN ISLAND
CITY UNIVERSITY OF NEW YORK
2800 VICTORY BOULEVARD
STATEN ISLAND, NY 10314
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 47–62

ON ADDITIVE FUNCTIONALS OF DIFFUSION PROCESSES

E. CSÁKI[1] and P. SALMINEN[2]

Dedicated to Pál Révész for his 60th birthday

Abstract

We study the properties of the local time and its inverse of a linear recurrent diffusion and establish strong approximation results for additive functionals of the diffusion. The cases of positive, resp. null recurrence are treated separately. A consequence of our results is a law of the iterated logarithm for additive functional.

1. Introduction

Let $X = \{X_t;\ t \geqq 0\}$ be a recurrent one-dimensional diffusion living on an interval $I \subseteq \Re$. Our aim in this paper is to extend the results of Csáki et al. [4] concerning strong approximations of additive functionals of the form

$$(1.1) \qquad Z_t = \int_0^t f(X_s)\, ds$$

established in the case when X was a standard Brownian motion.

Weak limit theorems for additive functionals of diffusions were proved by Tanaka [15] in the positive recurrent case and by Kasahara and Kotani [9] and Kasahara [8] in the null recurrent case.

For similar results for Markov chains see Csáki and Csörgő [2].

Horváth and Khoshnevisan [6] proved a strong approximation result for additive functionals of Ornstein–Uhlenbeck process.

Our results are based on the properties of the local time and its inverse which will be summarized in Section 2. These results may be of independent interest. In Section 3 we formulate and prove our main results in both positive and null recurrent cases. In Section 4 we give some remarks.

1991 *Mathematics Subject Classification*. Primary 60J60; Secondary 60F17, 60J55.
Key words and phrases. Diffusion, local time, additive functional.

[1] Research supported by Hungarian National Foundation for Scientific Research Grant No. 1905.

[2] Research supported in part by the Academy of Finland and the Hungarian Academy of Sciences on the basis of the scientific exchange agreement between Hungary and Finland.

2. Diffusion local times

Let $X = \{X_t,\ t \geqq 0\}$ be a recurrent one-dimensional diffusion living on an interval $I \subseteq \Re$. Recall that a recurrent diffusion is null recurrent if and only if $m(I) = \infty$, where m is the speed measure of X. If $m(I) < \infty$ we say that X is positively recurrent.

The probability measure and the expectation operator of X, when $X_0 = x$, are denoted by $\mathbf{P}_x$ and $\mathbf{E}_x$, respectively. The diffusion X possesses at every point $x \in I$ a local time process $L^x = \{L^x_t,\ t \geqq 0\}$, such that (see Ray [13])

$$L^x_t = \lim_{\epsilon \downarrow 0} \frac{1}{m((x-\epsilon, x+\epsilon))} \int_0^t 1_{(x-\epsilon, x+\epsilon)}(X_s)\,ds \qquad \mathbf{P}-\text{a.s.}, \tag{2.1}$$

$$(t,x) \mapsto L^x_t \quad \text{is continuous } \mathbf{P}-\text{a.s.} \tag{2.2}$$

Let now A^x be the right continuous inverse of L^x. Then it is well-known that A^x is a subordinator, i.e. an increasing process with stationary and independent increments. The Lévy–Khintchine representation of A^x is (see Itô and McKean [7], p. 214–215)

$$\begin{aligned} \mathbf{E}_x(e^{-\lambda A^x_t}) &= \exp\left(-t\int_0^\infty (1-e^{-\lambda l})n^x(dl)\right) \\ &= \exp\left(-\frac{t}{G_\lambda(x,x)}\right), \end{aligned} \tag{2.3}$$

where G_λ is the Green kernel of X. The Lévy measure n^x is the sum of the two measures n^x_+ and n^x_- obtained from the relations

$$\int_0^\infty (1-e^{-\lambda l})n^x_\pm(dl) = \lim_{y\to x\pm} \frac{1-\mathbf{E}_y(\exp(-\lambda H_x))}{\pm(S(y)-S(x))}, \tag{2.4}$$

where S is the scale function of X and $H_x = \inf\{t,\ X_t = x\}$. An implicit assumption here is that x is an inner point of I. If x is the left endpoint, say, then by the definition $n_- \equiv 0$. We remark also that

$$\mathbf{E}_x(A^x_1) = \lim_{\lambda\to 0} \frac{1}{\lambda G_\lambda(x,x)} = m\{I\} \tag{2.5}$$

(see Salminen [14]).

Further, let $B = \{B_s, s \geqq 0\}$ be a standard one-dimensional Brownian motion and $l = \{l^x_t, t \geqq 0, x \in \Re\}$ its local time (with respect to $2dx$, where

dx is the Lebesgue measure). The inverse local time of B at the point x is denoted by $a^x = \{a_t^x, t \geqq 0\}$.

It is a standard fact that X can be constructed from B via a scale transformation and a random time change. To explain shortly how this is done recall that X is said to be in natural scale if $S(x) = x$. Because S is increasing it is obvious that $Y = S(X)$ is a recurrent diffusion in natural scale. Let m^Y be the speed measure of Y, and introduce the additive functional

$$\alpha_t = \int_{\Re} l_t^y m^Y(dy). \tag{2.6}$$

Then $Y = \{Y_t, t \geqq 0\}$ is identical in law with $\{B(\tau_t), t \geqq 0\}$, where $\tau_t = \inf\{s, \alpha_s > t\}$ is the right continuous inverse of α (see Itô and McKean [7], p. 167–174).

Assume now that X is in natural scale and that $0 \in I$.

Proposition 2.1. *In a suitable probability space*

$$L_{A_t}^x = l_{a_t}^x \quad \text{for every } x \in I \text{ and } t \geqq 0 \quad a.s., \tag{2.7}$$

where $A_t = A_t^0$ and $a_t = a_t^0$.

Proof. First, because

$$A_t = \inf\{s,\ L_s^0 > t\} = \inf\{s,\ l_{\tau_s}^0 > t\},$$

we note that

$$\tau_{A_t} = \inf\{s,\ l_s^0 > t\} = a_t.$$

Using this and the definition of local time we have

$$\begin{aligned}
L_{A_t}^x &= \lim_{\epsilon \downarrow 0} \frac{1}{m((x-\epsilon, x+\epsilon))} \int_0^{A_t} 1_{(x-\epsilon,x+\epsilon)}(X_s)ds \\
&= \lim_{\epsilon \downarrow 0} \frac{1}{m((x-\epsilon, x+\epsilon))} \int_0^{A_t} 1_{(x-\epsilon,x+\epsilon)}(B(\tau_s))ds \\
&= \lim_{\epsilon \downarrow 0} \frac{1}{m((x-\epsilon, x+\epsilon))} \int_0^{a_t} 1_{(x-\epsilon,x+\epsilon)}(B_s)d\alpha_s \\
&= \lim_{\epsilon \downarrow 0} \frac{1}{m((x-\epsilon, x+\epsilon))} \int_{\Re} \int_0^{a_t} 1_{(x-\epsilon,x+\epsilon)}(y) dl_s^y m(dy)
\end{aligned}$$

$$= \lim_{\epsilon \downarrow 0} \frac{1}{m((x-\epsilon, x+\epsilon))} \int_{x-\epsilon}^{x+\epsilon} \int_0^{a_t} 1_{(x-\epsilon, x+\epsilon)}(y) dl_s^y m(dy)$$

$$= \lim_{\epsilon \downarrow 0} \frac{1}{m((x-\epsilon, x+\epsilon))} \int_{x-\epsilon}^{x+\epsilon} l_{a_t}^y m(dy)$$

$$= l_{a_t}^x,$$

by the continuity of $x \mapsto l_{a_t}^x$. The proposition follows using the joint continuity (2.2).

PROPOSITION 2.2. *For every $x \in I$ and $t > 0$*

$$\mathbf{E}_0^X(e^{-\beta L_{A_t}^x}) = \mathbf{E}_0^B(e^{-\beta l_{a_t}^x}) = \exp\left(-\frac{\beta t}{1+\beta|x|}\right). \tag{2.8}$$

Moreover, under $\mathbf{P}_0^X$ and for $x < 0 < y$ the random variables $L_{A_t}^x$ and $L_{A_t}^y$ are independent, and $L_{A_t} = \{L_{A_t}^x,\ x \geqq 0\}$ is a diffusion starting from $t > 0$ having the generator

$$\mathcal{G} = y\frac{d^2}{dx^2}. \tag{2.9}$$

Hence, L_{A_t} constitutes a martingale with respect to its natural filtration.

PROOF. This follows easily from Proposition 2.1 and the corresponding properties of Brownian local time (see Theorem 5.3.20, p. 137 in Knight [10]). (In fact, our case can also be seen as a special case of Ray [13], p. 624 and 627 by choosing therein $T = \inf\{s, L_s^0 > \tau\}$, where $\tau \sim \exp(\lambda)$.)

REMARK. Let $0 < x_1 < \dots < x_n$ be given. Then the joint distribution of

$$(L_{A_t}^{x_1}, \dots, L_{A_t}^{x_n})$$

is given by

$$\mathbf{E}_0(e^{-\sum \lambda_i L_{A_\tau}^{x_i}}) = \lambda G_0(0,0), \tag{2.9}$$

where $\tau \sim \exp(\lambda)$ and G_0 is the Green kernel associated with the diffusion having the generator

$$\begin{aligned} \mathcal{G}u &= \frac{d}{dm}\frac{d}{dx}u, \\ u'(0+) - u'(0-) &= \lambda u(0) \\ u'(x_i+) - u'(x_i-) &= \lambda_i u(x_i), \quad i = 1, 2, \dots, n. \end{aligned} \tag{2.11}$$

PROPOSITION 2.3. *Let X be in natural scale. Then the following formulae are valid for $x \in I$*

$$\mathbf{E}_0(L^x_{A_t}) = t \tag{2.12}$$

$$\mathbf{E}_0((L^x_{A_t})^2) = t^2 + 2t|x| \tag{2.13}$$

$$\mathbf{E}_0((L^x_{A_1})^m) \leqq K_m(1 + |x|^{m-1}), \quad m \geqq 1 \tag{2.14}$$

$$\operatorname{Cov}(L^x_{A_t}, L^y_{A_t}) = \begin{cases} 2t \min(|x|, |y|), & \text{if } xy > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{2.15}$$

PROOF is straightforward and omitted (cf. Csáki et al. [4]).

Assume now that X is not in natural scale and that $0 \in I$ and — without loss of generality — $S(0) = 0$. Then $Y = S(X)$ is in natural scale and we have

$$\mathbf{E}^Y_a \left(\exp(-\beta L^b_{A^Y_t})\right) = \exp\left(-\frac{\beta t}{1 + \beta|b - a|}\right), \tag{2.16}$$

where $A^Y_t = \inf\{s, L^a_s > t\}$. Consequently,

$$\mathbf{E}^X_{S^{-1}(a)} \left(\exp(-\beta L^{S^{-1}(b)}_{A^X_t})\right) = \exp\left(-\frac{\beta t}{1 + \beta|b - a|}\right), \tag{2.17}$$

where $A^X_t = \inf\{s, L^{S^{-1}(a)}_s > t\}$. Let now $a = S(0) = 0$ and $b = S(x)$, then

$$\mathbf{E}^X_0 (e^{-\beta L^x_{A_t}}) = \exp\left(-\frac{\beta t}{1 + \beta|S(x)|}\right). \tag{2.18}$$

The following proposition can easily been proved using Proposition 2.2 and the scale transformation.

PROPOSITION 2.4. *The following formulae are valid for $x \in I$*

$$\mathbf{E}_0(L^x_{A_t}) = t \tag{2.19}$$

$$\mathbf{E}_0((L^x_{A_t})^2) = t^2 + 2t|S(x)| \tag{2.20}$$

$$\mathbf{E}_0((L^x_{A_1})^m) \leqq K_m(1 + |S(x)|^{m-1}), \quad m \geqq 1 \tag{2.21}$$

$$\operatorname{Cov}(L^x_{A_t}, L^y_{A_t}) = \begin{cases} 2t \min(|S(x)|, |S(y)|), & \text{if } xy > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{2.22}$$

3. Main results

3.1 Preliminaries. Assume that $f(x)$, $x \in I$ is a locally integrable real valued function with the property

$$\int_I |f(x)| m(dx) < \infty. \tag{3.1}$$

Assume throughout that $X_0 = 0$ and consider the additive functional

$$Z_t = \int_0^t f(X_s)\,ds = \int_I f(x) L_t^x m(dx) \tag{3.2}$$

and put $t = A_u$:

$$Z_{A_u} = \int_0^{A_u} f(X_s)\,ds = \int_I f(x) L_{A_u}^x \, m(dx). \tag{3.3}$$

It is easy to see that $\{Z_{A_u},\, u \geqq 0\}$ is a process of independent and stationary increments. By (2.19) and (2.22) we have

$$\mathbf{E} Z_{A_u} = u\bar{f}, \tag{3.4}$$

where

$$\bar{f} = \int_I f(x)\, m(dx). \tag{3.5}$$

and

$$\operatorname{Var} Z_{A_u} = u\sigma^2, \tag{3.6}$$

where

$$\sigma^2 = 2 \iint\limits_{I \times I \cap \{xy>0\}} f(x) f(y) \min(|S(x)|, |S(y)|)\, m(dx) m(dy). \tag{3.7}$$

3.2 The case of positive recurrence. In this case, by (2.5),

$$\mu = \mathbf{E} A_1 = m\{I\} \tag{3.8}$$

is finite.

THEOREM 3.1. *Assume that*

$$\mathbf{E}(A_1)^q < \infty \quad \textit{for some} \quad 1 < q \leqq 2. \tag{3.9}$$

(i) *If*

$$\mathbf{E}\left(\int_0^{A_1} |f(X_s)|\,ds\right)^{2+\delta} < \infty \tag{3.10}$$

for some $\delta > 0$, then on a suitable probability space one can construct a diffusion process X_t and a standard Brownian motion W_t such that

$$\left| Z_t - \bar{f} L_t^0 - \frac{\sigma}{\sqrt{\mu}} W_t \right| = O(t^\beta \log t) \quad a.s. \tag{3.11}$$

as $t \to \infty$, where

$$\beta = \max\left(\frac{1}{2+\delta}, \frac{1}{2q} \right) \tag{3.12}$$

and σ^2 is defined by (3.7).

(ii) *If*

$$\mathbf{E} \left(\int_0^{A_1} \left| f(X_s) - \frac{\bar{f}}{\mu} \right| ds \right)^{2+\delta} < \infty \tag{3.13}$$

for some $\delta > 0$, then on a suitable probability space one can construct a diffusion process X_t and a standard Brownian motion W_t such that

$$\left| Z_t - \bar{f} \frac{t}{\mu} - \frac{\sigma_1}{\sqrt{\mu}} W_t \right| = O(t^\beta \log t) \quad a.s. \tag{3.14}$$

as $t \to \infty$, where β is defined by (3.12) *and*

(3.15)

$$\sigma_1^2 = 2 \iint_{I \times I \cap \{xy > 0\}} \left(f(x) - \frac{\bar{f}}{\mu} \right) \left(f(y) - \frac{\bar{f}}{\mu} \right) \min(|S(x)|, |S(y)|)\, m(dx) m(dy).$$

PROOF. We prove part (i) only. The proof of part (ii) is similar. It follows from the strong invariance principle of Komlós et al. [11] combined with Lemma A1 in Berkes and Philipp [1] that on a suitable probability space one can construct a diffusion process X_t and a standard Brownian motion $W_u^{(1)}$ such that

$$\left| Z_{A_u} - \bar{f} u - \sigma W_u^{(1)} \right| = O(u^{1/(2+\delta)}) \quad \text{a.s.} \tag{3.16}$$

as $u \to \infty$.

Under the condition (3.6) we have

$$A_u = \mu u + O(u^{1/q} (\log u)^{1/2}) \qquad \text{a.s.} \tag{3.17}$$

as $u \to \infty$ and consequently

$$L_t^0 = \frac{t}{\mu} + O(t^{1/q} (\log t)^{1/2}) \qquad \text{a.s.} \tag{3.18}$$

as $t \to \infty$. Put $u = L_t^0$ into (3.16). Then we have

$$\left| Z_{A_{L_t^0}} - fL_t^0 - \sigma W_{L_t^0}^{(1)} \right| = O\left((L_t^0)^{1/(2+\delta)}\right) = O(t^{1/(2+\delta)}) \qquad \text{a.s.} \tag{3.19}$$

as $t \to \infty$.

Under the condition (3.10) it follows that $\int_{A_k}^{A_{k+1}} |f(X_s)|\, ds$, $k = 1, 2, \dots,$ are i.i.d. random variables having $(2+\delta)$-th moment, therefore

$$|Z_{A_{L_t^0}} - Z_t| = O((L_t^0)^{1/(2+\delta)}) = O(t^{1/(2+\delta)}) \qquad \text{a.s.} \tag{3.20}$$

as $t \to \infty$. Moreover, it follows from (3.18) and the increment results of Csörgő and Révész [5] that

$$|W_{L_t^0}^{(1)} - W_{t/\mu}^{(1)}| = O(t^{1/(2q)} \log t) \qquad \text{a.s.} \tag{3.21}$$

as $t \to \infty$. Now our statement (3.11) follows from (3.19), (3.20), (3.21) and the scale change property of Brownian motion, i.e. $W_t = \sqrt{\mu} W_{t/\mu}^{(1)}$ is also a Brownian motion.

3.3 The case of null recurrence. In this case $m\{I\}$ and hence $\mathbf{E}A_1$ is infinite. We assume in fact, that A_u is close to a stable process T_u of order α, more precisely, assume that on a suitable probability space one can construct a diffusion X and a stable process T of order α such that

$$\mathbf{E}(e^{-\lambda T_u}) = e^{-cu\lambda^\alpha}, \quad \lambda > 0$$

with some positive constant c and

$$|A_u - T_u| = O(u^\kappa) \qquad \text{a.s.} \tag{3.22}$$

as $u \to \infty$ for some $0 \le \kappa < 1/\alpha$.

On the other hand, it follows from the invariance principle of Komlós et al. [11] combined with Lemma A1 in Berkes and Philipp [1] that under the condition

$$\mathbf{E}|Z_{A_1}|^{2+\delta} < \infty, \tag{3.23}$$

which is a consequence of (3.10), on a suitable probability space one can construct a diffusion X and a standard Brownian motion W such that

$$|Z_{A_u} - fu - \sigma W_u| = O(u^{1/(2+\delta)}) \quad \text{a.s.} \tag{3.24}$$

as $u \to \infty$, where f is defined by (3.5) and σ is defined by (3.7).

Now we state our main result in the null recurrent case.

THEOREM 3.2. *Assume that X is a null recurrent diffusion process on an interval I, $0 \in I$ with local time L_t^x and that* (3.22) *and* (3.10) *hold true. Then on a suitable probability space one can construct the diffusion X_t, a standard Brownian motion W_u and a nondecreasing stable process T_u of order α such that W and T are independent and for $\varepsilon > 0$ small enough we have*

$$|Z_t - \bar{f} L_t^0 - \sigma W_{V_t}| = O(t^{\alpha/2-\varepsilon}) \quad a.s. \tag{3.25}$$

as $t \to \infty$, where V_t is the (continuous) inverse of T_u. Z_t, $\bar{f}$ and σ, resp. are defined by (3.2), (3.5) *and* (3.7), *resp.*

PROOF goes along the same line as that of Theorem 2 in Csáki et al. [4] and Theorem 3.1 in Csáki and Csörgő [2]. In order to assure independence of W and T, start from two independent copies, $X^{(1)}$ and $X^{(2)}$ of the diffusion process X and use the approximation (3.22) to $X^{(1)}$ and the approximation (3.24) to $X^{(2)}$, i.e. assume that we have with independent W and T

$$|A_u^{(1)} - T_u| = O(u^{\kappa}) \qquad \text{a.s.} \tag{3.26}$$

and

$$|Z^{(2)}_{A_u^{(2)}} - \bar{f} u - \sigma W_u| = O(u^{1/(2+\delta)}) \quad \text{a.s.} \tag{3.27}$$

as $u \to \infty$, where the superscripts (1) and (2), resp. indicate that the corresponding quantities are defined in terms of $X^{(1)}$ and $X^{(2)}$, resp. Now define a new diffusion process X_t as follows: put $u_{-1} = 0$, $u_k = 2^k$, $r_k = u_k - u_{k-1} = 2^{k-1}$, $k = 1, 2, \ldots$ and consider the k-th block as

$$\{X^{(j)}_{A_u}, \ u_{k-1} \leqq u < u_k\}, \quad j = 1, 2. \tag{3.28}$$

The section between $A^{(j)}_{u_{k-1}+i-1}$ and $A^{(j)}_{u_{k-1}+i}$, $i = 1, 2, \ldots, r_k$ will be called excursion. Call an excursion large in the k-th block if

$$A^{(j)}_{u_{k-1}+i} - A^{(j)}_{u_{k-1}+i-1} > r_k^{\theta} \tag{3.29}$$

and small in the k-th block if

$$A^{(j)}_{u_{k-1}+i} - A^{(j)}_{u_{k-1}+i-1} \leqq r_k^{\theta}, \tag{3.30}$$

where

$$\theta = \frac{1}{\alpha} - \frac{1}{2} + \frac{1}{2+\delta}. \tag{3.31}$$

Now construct the new diffusion X from the large excursions of $X^{(1)}$ and small excursions of $X^{(2)}$ as described in Csáki et al. [4]. The quantities defined in terms of the new diffusion will be denoted without superscripts. Then as in Csáki et al. [4] or in Csáki and Csörgő [2], we have

$$(3.32)\qquad \begin{aligned} F_k &:= \max_{1\leqq i\leqq r_k} |A_{u_{k-1}+i} - A_{u_{k-1}} - (A^{(1)}_{u_{k-1}+i} - A^{(1)}_{u_{k-1}})| \leqq \\ &\leqq H_k := \sum_{j=1}^{2}\sum_{i=1}^{r_k}(A^{(j)}_{u_{k-1}+i} - A^{(j)}_{u_{k-1}+i-1}) I\{A^{(j)}_{u_{k-1}+i} - A^{(j)}_{u_{k-1}+i-1} \leqq r_k^{\theta}\} \end{aligned}$$

and denoting $Y_u = Z_{A_u}$,

$$(3.33)\qquad \begin{aligned} G_k &:= \max_{1\leqq i\leqq r_k} |Y_{u_{k-1}+i} - Y_{u_{k-1}} - (Y^{(2)}_{u_{k-1}+i} - Y^{(2)}_{u_{k-1}})| \leqq \\ &\leqq J_k := 2(\nu_k^{(1)} + \nu_k^{(2)}) \max_{j=1,2}\max_{1\leqq i\leqq r_k} |Y^{(j)}_{u_{k-1}+i} - Y^{(j)}_{u_{k-1}+i-1}|, \end{aligned}$$

where $\nu_k^{(j)}$ denotes the number of large excursions in the k-th block.

It follows from (3.22) that the random variable A_1 is in the normal domain of attraction of the stable law of order α, hence with some constant c_1,

$$(3.34)\qquad \mathbf{P}(A_1 \geqq v) \sim c_1 v^{-\alpha}, \quad v \to \infty$$

and therefore

$$(3.35)\qquad \mathbf{E}(A_1 I\{A_1 \leqq z\}) = \int_0^z \mathbf{P}(A_1 \geqq v)\,dv \leqq c_2 z^{1-\alpha},$$

with some constant c_2, consequently from (3.32)

$$(3.36)\qquad \mathbf{E}F_k \leqq \mathbf{E}H_k \leqq 2c_2 r_k^{1+\theta(1-\alpha)},$$

thus by Markov's inequality

$$(3.37)\qquad \mathbf{P}(F_k \geqq 2^{\kappa(k-1)}) \leqq 2c_2 2^{(k-1)(1+\theta(1-\alpha)-\kappa)}.$$

We may assume without loss of generality that $1+\theta(1-\alpha) < \kappa$ and so by Borel-Cantelli lemma

$$(3.38)\qquad F_k = O(2^{\kappa(k-1)}) \qquad \text{a.s.}$$

and hence

$$(3.39)\qquad \max_{i\leqq u_k} |A_i - A_i^{(1)}| = O(2^{\kappa k}) \qquad \text{a.s.}$$

as $k \to \infty$ and

$$(3.40)\qquad |A_u - A_u^{(1)}| = O(u^{\kappa}) \qquad \text{a.s.}$$

as $u \to \infty$.

On the other hand, $\nu_k^{(1)} + \nu_k^{(2)}$ has a binomial distribution with parameters $2r_k$ and $p = \mathbf{P}(A_1 \geqq r_k^\theta) \leqq c_3 r_k^{-\theta\alpha}$. We get by Markov's inequality

$$\begin{aligned}\mathbf{P}(\nu_k^{(1)} + \nu_k^{(2)} > 4c_3 r_k^{1-\theta\alpha}) &\leqq \mathbf{P}(e^{\nu_k^{(1)}+\nu_k^{(2)}} > e^{4c_3 r_k^{1-\theta\alpha}}) \\ &\leqq (1+p(e-1))^{2r_k} e^{-4c_3 r_k^{1-\theta\alpha}} \leqq e^{2c_3(e-3)r_k^{1-\theta\alpha}}.\end{aligned}$$

Hence

$$\begin{aligned}\mathbf{P}(G_k \geqq r_k^\tau) &\leqq \mathbf{P}(J_k \geqq r_k^\tau) \leqq \mathbf{P}(\nu_k^{(1)} + \nu_k^{(2)} \geqq 4c_3 r_k^{1-\theta\alpha}) + \\ &+ \mathbf{P}(8c_3 r_k^{1-\theta\alpha} \max_{j=1,2} \max_{1 \leqq i \leqq r_k} |Y^{(j)}_{u_{k-1}+i} - Y^{(j)}_{u_{k-1}+i-1}| \geqq r_k^\tau) \leqq \\ &\leqq \exp\{2c_3(e-3) r_k^{1-\theta\alpha}\} + \mathbf{P}(8c_3 \max_{j=1,2} \max_{1 \leqq i \leqq r_k} |Y^{(j)}_{u_{k-1}+i} - Y^{(j)}_{u_{k-1}+i-1}| \geqq r_k^{\tau-1+\theta\alpha}) \\ &\leqq \exp\{2c_3(e-3) r_k^{1-\theta\alpha}\} + 2r_k \mathbf{P}(8c_3 |Z_{A_1}| \geqq r_k^{\tau-1+\theta\alpha}) \leqq \\ &\leqq \exp\{2c_3(e-3) r_k^{1-\theta\alpha}\} + 2(8c_3)^{2+\delta} \mathbf{E}|Z_{A_1}|^{2+\delta} r_k^{1-(\tau-1+\theta\alpha)(2+\delta)} \leqq \\ &\leqq C r_k^{1-(\tau-1+\theta\alpha)(2+\delta)},\end{aligned}$$

where θ is defined by (3.31).

Now choosing τ so that

(3.41) $$\frac{1}{2+\delta} + \frac{\alpha\delta}{2(2+\delta)} < \tau < \frac{1}{2},$$

we get by Borel-Cantelli lemma

(3.42) $$G_k = O(2^{(k-1)\tau}) \qquad \text{a.s.}$$

as $k \to \infty$, or

(3.43) $$|Z_{A_u} - Z^{(2)}_{A_u^{(2)}}| = O(u^\tau) \qquad \text{a.s.}$$

as $u \to \infty$. By (3.26), (3.27), (3.40) and (3.43) we have

(3.44) $$|Z_{A_u} - fu - \sigma W_u| = O(u^\tau) \qquad \text{a.s.}$$

and

(3.45) $$|A_u - T_u| = O(u^\kappa) \qquad \text{a.s.}$$

as $u \to \infty$.

Now put $u = L_t^0$ into (3.44):

(3.46) $$|Z_{A_{L_t^0}} - fL_t^0 - \sigma W_{L_t^0}| = O\left((L_t^0)^\tau\right) \qquad \text{a.s.}$$

as $t \to \infty$.

For the stable process T we have for any $\varepsilon > 0$ and large enough u (cf. Mijnheer [12])

$$u^{1/\alpha-\varepsilon} \leqq T_u \leqq u^{1/\alpha+\varepsilon} \qquad \text{a.s.} \tag{3.47}$$

and also by (3.45)

$$u^{1/\alpha-\varepsilon} \leqq A_u \leqq u^{1/\alpha+\varepsilon} \qquad \text{a.s.} \tag{3.48}$$

implying

$$t^{\alpha-\varepsilon} \leqq V_t \leqq t^{\alpha+\varepsilon} \qquad \text{a.s.} \tag{3.49}$$

and

$$t^{\alpha-\varepsilon} \leqq L_t^0 \leqq t^{\alpha+\varepsilon} \qquad \text{a.s.} \tag{3.50}$$

for any $\varepsilon > 0$ and large enough t.

Next we show that for any $\varepsilon > 0$,

$$|V_t - L_t^0| = O(t^{\alpha^2\kappa+\varepsilon}) \qquad \text{a.s.} \tag{3.51}$$

as $t \to \infty$.

We need the following increment results:

Lemma 3.1. *For $0 < \beta < 1$ and $\varepsilon > 0$ we have*

$$\sup_{0 \leqq s \leqq t} (V_{s+t^\beta} - V_s) = O(t^{\alpha\beta+\varepsilon}) \qquad \textit{a.s.} \tag{3.52}$$

and

$$\sup_{0 \leqq s \leqq t} (L^0_{s+t^\beta} - L^0_s) = O(t^{\alpha\beta+\varepsilon}) \qquad \textit{a.s.} \tag{3.53}$$

as $t \to \infty$.

Proof. For $s \geqq 0$, $t > 0$ we have

$$\begin{aligned} \mathbf{P}(V_{s+2t^\beta} - V_s > u) &\leqq \mathbf{P}(V_{2t^\beta} > u) \\ = \mathbf{P}(T_u \leqq 2t^\beta) &= \mathbf{P}(e^{-\lambda T_u} \geqq e^{-2\lambda t^\beta}) \\ \leqq e^{2\lambda t^\beta}\mathbf{E}(e^{-\lambda T_u}) &= e^{2\lambda t^\beta - cu\lambda^\alpha}. \end{aligned}$$

By choosing

$$\lambda = \left(\frac{cu\alpha}{2t^\beta}\right)^{\frac{1}{1-\alpha}}, \tag{3.54}$$

we obtain that

$$\mathbf{P}(V_{s+2t^\beta} - V_s > u) \leqq \exp\left\{-c'\left(\frac{u}{t^{\alpha\beta}}\right)^{\frac{1}{1-\alpha}}\right\} \tag{3.55}$$

with some positive constant c'. Put $t = k+1$, $s = j$ integers, $u = k^{\alpha\beta+\varepsilon}$, then

$$\mathbf{P}\left(\sup_{0\leqq j\leqq k+1}(V_{(j+2(k+1)^\beta)} - V_j) > k^{\alpha\beta+\varepsilon}\right) \leqq (k+2)\exp\left\{-c'k^{\varepsilon/(1-\alpha)}\right\},$$

which is summable, so by Borel–Cantelli lemma

$$\sup_{0\leqq j\leqq k+1}(V_{j+2(k+1)^\beta} - V_j) = O(k^{\alpha\beta+\varepsilon}) \qquad \text{a.s.}$$

as $k \to \infty$. Since for $k \leqq t < k+1$

$$\frac{\sup_{0\leqq s\leqq t}(V_{s+t^\beta} - V_s)}{t^{\alpha\beta+\varepsilon}} \leqq \frac{2\sup_{0\leqq j\leqq k+1}(V_{j+2(k+1)^\beta} - V_j)}{k^{\alpha\beta+\varepsilon}}, \tag{3.56}$$

we have also (3.52).

To show (3.53) we note that since A_1 is in the normal domain of attraction of the stable law of T_1,

$$\log \mathbf{E}(e^{-\lambda A_1}) \sim -c\lambda^\alpha, \quad \lambda \to 0 \tag{3.57}$$

and we can proceed as above. This proves the Lemma.

Now we are ready to prove (3.51). Consider first t-s for which $V_t < L_t^0$. Then $A_{L_t^0} \leqq t < T_{L_t^0}$, hence (3.45) and (3.50) together imply that

$$0 \leqq T_{L_t^0} - t \leqq T_{L_t^0} - A_{L_t^0} = O(t^{\kappa(\alpha+\varepsilon)}) \qquad \text{a.s.} \tag{3.58}$$

as $t \to \infty$. Since $L_t^0 = V_{T_{L_t^0}}$, it follows from (3.52) and (3.58) that

$$|L_t^0 - V_t| = |V_{T_{L_t^0}} - V_t| = O(t^{\kappa\alpha(\alpha+\varepsilon)+\varepsilon}) \qquad \text{a.s.} \tag{3.59}$$

as $t \to \infty$ on the set $\{t : V_t < L_t^0\}$. For $\{t : L_t^0 \leqq V_t\}$ we can prove (3.59) similarly, interchanging the role of V_t and L_t^0. Since $\varepsilon > 0$ is arbitrary, we obtain (3.51).

It follows from the increment results for Brownian motion due to Csörgő and Révész [5] that

$$|W_{L_t^0} - W_{V_t}| = O(t^{\alpha^2\kappa/2+\varepsilon}) \qquad \text{a.s.} \tag{3.60}$$

for any $\varepsilon > 0$, as $t \to \infty$.

One can see moreover that

$$|Z_{A_{L_t^0}} - Z_t| \leqq \int_{A_{L_t^0}}^{t} |f(X_s)|\, ds \leqq \int_{A_{[L_t^0]}}^{A_{[L_t^0]+1}} |f(X_s)|\, ds. \tag{3.61}$$

The random variables $\int_{A_k}^{A_{k+1}} |f(X_s)|\, ds$, $k = 1, 2, \dots$, are i.i.d. and have $(2+\delta)$-th moments by the condition (3.10), therefore

$$\int_{A_k}^{A_{k+1}} |f(X_s)|\, ds = O(k^{1/(2+\delta)}) \qquad \text{a.s.} \tag{3.62}$$

as $k \to \infty$. (3.50), (3.61) and (3.62) together imply that

$$|Z_{A_{L_t^0}} - Z_t| = O(t^{\alpha/(2+\delta)+\varepsilon}) \qquad \text{a.s.} \tag{3.63}$$

as $t \to \infty$.

Finally, (3.25) follows from (3.41), (3.46), (3.50), (3.60) and (3.63). This completes the proof of Theorem 3.2.

4. Some remarks

First some comments on our conditions:

(3.10) is satisfied by a broad class of functions f. E.g. in the case when X is in natural scale and is positive recurrent, then f being bounded is a sufficient condition for (3.10). In the null recurrent case if f is bounded and has compact support, then (3.10) is satisfied. Another fairly general sufficient condition for (3.10) is

$$\int_I |S(x)|^{1+\delta} |f(x)| m(dx) < \infty. \tag{4.1}$$

As far as the condition (3.22) is concerned, we note that if X is in natural scale, i.e. $S(x) = x$ and the speed measure is $m(dx) = 2|x|^{\beta} dx$, $\beta > -1$, then (cf. Itô and McKean [7], p. 226) A_u is a stable process of order $\alpha = 1/(\beta+2)$. So we can take $T_u = A_u$ in this case.

Our strong approximation results imply weak convergence of additive functionals. But our conditions are stronger than needed (cf. Tanaka [15], Kasahara and Kotani [9] and Kasahara [8]). The merit of the strong approximation is that we can conclude certain strong limit theorems. It follows e.g. from Theorem 3.1 that under its conditions we have the same strong limit theorems (Strassen's LIL, Chung's LIL, etc.) for $(Z_t - \bar{f} L_t^0)\sqrt{\mu}\sigma^{-1}$ in the case (i) and for $(Z_t - \bar{f}t/\mu)\sqrt{\mu}\sigma_1^{-1}$ in the case (ii) as for W_t. Concerning Theorem 3.2, we can obtain a law of the iterated logarithm for W_{V_t}. Write

$$\frac{W_{V_t}}{t^{\alpha/2}} = \frac{W_{V_t}}{\sqrt{V_t}}\sqrt{\frac{V_t}{t^\alpha}}. \tag{4.2}$$

Here $W_{V_t}/\sqrt{V_t}$ is a standard normal variable, independent of V_t. We know that T_1 has stable distribution of order α, so for its density we have (cf. Mijnheer [12])

$$\log p_{T_1}(x) \sim -\frac{1-\alpha}{\alpha}(c\alpha x^{-\alpha})^{1/(1-\alpha)}, \quad x \to 0 \tag{4.3}$$

and since V is the inverse of T, one can obtain for its density

$$\log p_{V_1}(x) \sim -\frac{1-\alpha}{\alpha}(c\alpha x)^{1/(1-\alpha)}, \quad x \to \infty. \tag{4.4}$$

V_t/t^α has the same distribution as V_1, so one can obtain for the density of $W_{V_t}/t^{\alpha/2}$ that

$$\log p_{W_{V_1}}(x) \sim -(2-\alpha)\left(\frac{\alpha^\alpha c}{2}\right)^{1/(2-\alpha)} x^{2/(2-\alpha)}, \quad x \to \infty. \tag{4.5}$$

From this asymptotic relation one can prove by standard method the following law of the iterated logarithm:

$$\limsup_{t\to\infty} \frac{W_{V_t}}{t^{\alpha/2}(\log\log t)^{1-\alpha/2}} = \sqrt{\frac{2}{c}}\frac{1}{\alpha^{\alpha/2}(2-\alpha)^{1-\alpha/2}} \quad \text{a.s.}. \tag{4.6}$$

A corollary to Theorem 3.2 is that under its conditions, the same LIL holds for the additive functional

$$\limsup_{t\to\infty} \frac{(Z_t - \bar{f} L_t^0)\sigma^{-1}}{t^{\alpha/2}(\log\log t)^{1-\alpha/2}} = \sqrt{\frac{2}{c}}\frac{1}{\alpha^{\alpha/2}(2-\alpha)^{1-\alpha/2}} \quad \text{a.s.}. \tag{4.7}$$

REFERENCES

[1] Berkes, I. and Philipp, W., Approximation theorems for independent and weakly dependent random vectors, *Ann. Probab.* **7** (1979), 29–54. *MR* **80k**:60008

[2] Csáki, E. and Csörgő, M., On additive functionals of Markov chains, *J. Theor. Probab.*

[3] Csáki, E., Csörgő, M., Földes, A. and Révész, P., Brownian local time approximated by a Wiener sheet, *Ann. Probab.* **17** (1989), 516–537. *MR* **90b**:60102

[4] CSÁKI, E., CSÖRGŐ, M., FÖLDES, A. and RÉVÉSZ, P., Strong approximation of additive functionals, *J. Theor. Probab.* **5** (1992), 679–706. *MR* **93k**:60073

[5] CSÖRGŐ, M. and RÉVÉSZ, P., *Strong approximations in probability and statistics*, Probability and mathematical statistics, Academic Press, New York, 1981. *MR* **84d**:60050

[6] HORVÁTH, L. and KHOSHNEVISAN, D., A strong approximation for logarithmic averages, *Studia Sci. Math. Hungar.* **31** (1996), 187–196.

[7] ITÔ, K. and MCKEAN, H.P., *Diffusion processes and their sample paths*, Die Grundlehren der mathematischen Wissenschaften, Bd. 125, Academic Press, New York; Springer-Verlag, Berlin–New York, 1965. *MR* **33** #8031

[8] KASAHARA, Y., Two limit theorems for occupation times of Markov processes, *Japan J. Math. (N. S.)* **7** (1981), 291–300. *MR* **85h**:60113

[9] KASAHARA, Y. and KOTANI, S., On limit processes for a class of additive functionals of recurrent diffusion processes, *Z. Wahrsch. Verw. Gebiete* **49** (1979), 133–153. *MR* **80i**:60107

[10] KNIGHT, F.B., *Essentials of Brownian motion and diffusion*, Mathematical Surveys, 18, American Mathematical Society, Providence, R.I., 1981. *MR* **82m**:60098

[11] KOMLÓS, J., MAJOR, P. and TUSNÁDY, G., An approximation of partial sums of independent r.v.'s and the sample df. I, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **32** (1975), 111–131. *MR* **51** #11605b

[12] MIJNHEER, J.L., *Sample path properties of stable processes*, Mathematical Centre Tracts, No. 59, Mathematisch Centrum, Amsterdam, 1975. *MR* **51** #7008

[13] RAY, D., Sojourn times of diffusion processes, *Illinois J. Math.* **7** (1963), 615–630. *MR* **27** #6306

[14] SALMINEN, P., On the distribution of supremum of diffusion local time, *Statist. Probab. Lett.* **18** (1994), 219–225.

[15] TANAKA, H., Certain limit theorems concerning one-dimensional diffusion processes, *Mem. Fac. Sci. Kyusyu Univ. Ser. A Math.* **12** (1958), 1–11. *MR* **20** #3608

(Received November 24, 1994)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

MATHEMATICAL INSTITUTE
ÅBO AKADEMI
SF–20500 ÅBO
FINLAND

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 63–92

INVARIANCE PRINCIPLES IN BANACH FUNCTION SPACES

M. CSÖRGŐ[1] and R. NORVAIŠA[2]

Dedicated to P. Révész for his sixtieth birthday

Summary

We consider, as a sample path space, a Banach function space $\mathbb{F}$ of measurable functions defined on a σ-finite measure space. We pose the question of characterizing those Banach function spaces $\mathbb{F}$ where invariance principles hold true for distributions induced by stochastic processes with sample paths in $\mathbb{F}$. In this paper an answer to this question is given for Banach function spaces which do not contain l_∞^n's uniformly, for distributions induced by empirical type, partial sum and Poisson processes. We also give estimates of the rate of convergence in Prohorov metric between induced distributions. Proofs are based on strong approximation results of these processes and on describing their approximating Gaussian distributions on Banach function spaces.

1. Introduction

Some asymptotic results for empirical type and partial sum processes in a weighted sup-norm have been recently extended to L_p-norms (cf. the paragraph right after Theorem 2.1). The conditions for these results to hold true are different. An attempt to understand the nature of this difference motivated us to replace the L_p-norms by an arbitrary function norm on a space of measurable functions. For this we used strong approximation techniques in conjunction with some tools developed in Probability in Banach spaces. As a result we found that, in case of empirical processes the class of Banach function spaces for our invariance principles in probability to hold true, as well as for the thus resulting weak convergence statements, is larger than it would be if one were to use only the central limit theorem for Banach space valued random variables. As far as partial sum processes are concerned, the use of the general central limit theorem for triangular arrays of independent

1991 *Mathematics Subject Classification.* Primary 60F17, 60B11.

Key words and phrases. Empirical process, Brownian bridge, Poisson bridge, partial sum process, Wiener process, Poisson process, invariance principle, rate of convergence, Banach function space.

[1] Research supported by an NSERC Canada Grant at Carleton University, Ottawa, Canada.

[2] Research supported by an NSERC Canada Scientific Exchange Award at Carleton University, Ottawa, Canada.

0081-6906/96/$ 5.00 ©1996 *Akadémiai Kiadó, Budapest*

Banach space valued random variables is not straightforward at all because rows of the triangular arrays induced by representing partial sum processes this way are infinite. Indeed, the first question in this situation would be investigating the convergence of this series representation of a partial sum process in the Banach function space $\mathbb{F}$. This, however, is a special case of the open problem of giving necessary and sufficient conditions for the almost sure convergence in $\mathbb{F}$ of a series of independent $\mathbb{F}$-valued random variables. We note that, as a byproduct, a sufficient condition for this convergence problem also follows from our strong approximation approach. Moreover, we give estimates of the Prohorov metric as well in our invariance principles which are either best possible, and can be proved via Banach space techniques only under additional assumptions, or they are the first available such results. Thus by combining two techniques, we demonstrate a more powerful tool for solving the problems at hand than any one of them proved to be separately so far.

A brief description of the paper is as follows. We continue the introduction by recalling a weighted sup-norm version of approximation for the uniform empirical process, the Chibisov–O'Reilly theorem, in a form which will demonstrate the nature of the above mentioned difference. Section 2 contains the formulation and discussion of our results. The proofs are carried out in Section 5. The necessary facts about Banach function spaces to be used later on are given in Section 3 and the sample paths results for Gaussian and Poisson processes are collected in Section 4.

Let $U, U_1, U_2, \ldots$ be a sequence of independent uniform on $(0,1)$ random variables (rv's). For any integer $n \geqq 1$, define the *uniform empirical process* $E_n = \{E_n(t);\, t \in (0,1)\}$, based on a sample $U_1, \ldots, U_n$, by

$$E_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (1_{(0,t]}(U_i) - t), \qquad t \in (0,1),$$

where 1_A is the indicator function on $(0,1)$ of a set A. The order statistics of a sample $U_1, \ldots, U_n$ will be denoted by $U_{n:1} \leqq \cdots \leqq U_{n:n}$.

Let w be a positive function on $(0,1)$ which is nonincreasing in a neighbourhood of 0, is nondecreasing in a neighbourhood of 1 and is bounded on every compact subset of $(0,1)$. We call any such function a *weight* function on $(0,1)$. Given a weight function w on $(0,1)$, define a weighted sup-norm $\|\cdot\|_w \colon \Re^{(0,1)} \to [0,+\infty]$ by

$$\|f\|_w := \sup_{t \in (0,1)} w(t)|f(t)|$$

and define a vector space B_w by

$$B_w := \{\, f \in \Re^{(0,1)} \colon \|f\|_w < +\infty \,\}.$$

Endowed with the weighted sup-norm $\|\cdot\|_w$, B_w is a non-separable Banach space.

Now we are ready to formulate the Chibisov–O'Reilly theorem:

THEOREM A. *Let w be a weight function on $(0,1)$. The following statements about w are equivalent:*

(1) *w is a Chibisov–O'Reilly function, i.e.*

$$I(w,c) := \int_0^1 (t(1-t))^{-1} \exp\{-c(t(1-t))^{-1} w^{-2}(t)\} dt < +\infty,$$

for all $c > 0$;

(2) *one can construct a sequence of uniform empirical processes $\{E_n;\, n \geqq 1\}$ and a sequence of Brownian bridges $\{B_n;\, n \geqq 1\}$ in such a way that*

$$\|E_n - B_n\|_w = o_P(1);$$

(3) *any sequence of uniform empirical processes $\{E_n;\, n \geqq 1\}$ induce B_w-valued random functions which converge in law to a Radon probability γ on B_w, i.e. for every bounded continuous real valued function Φ on B_w we have*

$$\lim_{n\to\infty} E^*\Phi(E_n) = \int_{B_w} \Phi \, d\gamma,$$

where E^* denotes the upper integral.

The equivalence of (1) and (2) was proved by M. Csörgő, S. Csörgő, Horváth and Mason (cf. Theorem 4.2.1 in [7]) and that of (1) and (3) by Dudley (cf. Theorem 6.3 in [18]). The next statement is by combining of Theorems 4.2.2 and 4.2.3 in [7].

THEOREM B. *Let w be a weight function on $(0,1)$. The following statements about w are equivalent:*

(1) *w is a local function of a Brownian bridge, i.e. $I(w,c) < +\infty$ for some $c > 0$;*

(2) *one can construct a sequence of uniform empirical processes $\{E_n;\, n \geqq 1\}$ and a sequence of Brownian bridges $\{B_n;\, n \geqq 1\}$ in such a way that*

$$\|E_n - B_n\|_w = O_P(1);$$

(3) *for any sequence of uniform empirical processes $\{E_n;\, n \geqq 1\}$ we have*

$$\|E_n\|_w = O_P(1).$$

The following two remarks concern Theorem B.

REMARK 1. In [7] the class of weight functions which satisfy (1) of Theorem B was called the Erdős–Feller–Kolmogorov–Petrovski(EFKP) upper-class of a Brownian bridge B. However, Csörgő, Shao and Szyszkowicz pointed out in [17] that this class does not coincide with the classical definitions of upper- and lower-classes. Hence the weight functions satisfying (1) of Theorem B were simply called local functions.

REMARK 2. Statement (3) is called a *bounded central limit theorem* (cf., e.g., p. 276 in [24]). It follows from this statement that the sequence of rv's $\|E_n\|_w$ converges in distribution to a nondegenerate rv. By Theorem 4.2.3 in [7], the latter nondegenerate rv must be the rv $\|B\|_w$.

The implication (1) $\Rightarrow$ (2) of the next theorem follows from Theorem 2.2 of [7] (see also Proposition 4.1 below). For the proof of the converse implication (2) $\Rightarrow$ (1), we refer to the Appendix below.

THEOREM C. *Let w be a weight function on $(0,1)$. The following statements about w are equivalent:*

(1) $$\sup_{0<t<1} \sqrt{t(1-t)}w(t) < +\infty, \text{ i.e. the function } (0,1) \ni t \to \sqrt{t(1-t)} \in B_w;$$

(2) *one can construct a sequence of uniform empirical processes $\{E_n;\, n \geqq 1\}$ and a sequence of Brownian bridges $\{B_n;\, n \geqq 1\}$ in such a way that*

$$\|E_n - \bar{B}_n\|_w = O_P(1),$$

where

$$\bar{B}_n(t) = \begin{cases} B_n(t), & \text{if } t \in [U_{n:1}, U_{n:n}], \\ 0, & \text{elsewhere.} \end{cases}$$

2. Results

This section contains formulations and discussions of our results. Roughly speaking a function norm $\|\cdot\|$ is a norm on a subspace, called a Banach function space $\mathbb{F} = (\mathbb{F}(T,m), \|\cdot\|)$, of measurable functions defined on a σ-finite measure space (T,m). For a more precise definition and other notations we refer to Section 3.

Turning to our first result, it will be shown that condition (1) in Theorem C with the weighted sup-norm replaced by a function norm, characterizes the asymptotic results of Theorems A, B and C.

THEOREM 2.1. *Let $\mathbb{F} = (\mathbb{F}(T,m), \|\cdot\|)$, $T = (0,1)$, be a separable Banach function space which does not contain l_∞^n's uniformly. The following statements about $\mathbb{F}$ are equivalent:*

(1) *the function*

$$T \ni t \to \sqrt{t(1-t)} \in \mathbb{F}(T,m); \tag{2.1}$$

(2) *one can construct a sequence of uniform empirical processes $\{E_n;\, n \geqq 1\}$, and a sequence of Brownian bridges $\{B_n;\, n \geqq 1\}$ in such a way that E_n, as well as B_n, have almost all sample paths in $\mathbb{F}$ and*

$$\|E_n - B_n\| = o_P(1); \tag{2.2}$$

(3) *there exists a Brownian bridge distribution* $\mathcal{L}(B)$ *on* F *and a uniform empirical process* E_n *induces the distribution* $\mathcal{L}(E_n)$ *on* F *in such a way that*

$$\mathcal{L}(E_n) \Rightarrow \mathcal{L}(B) \quad \text{on} \quad \mathsf{F} \quad \text{weakly}\,; \tag{2.3}$$

(4) *for any uniform empirical process* E_n, *we have*

$$\|E_n\| = O_P(1).$$

The first results on the distributions of L_p norms of weighted uniform empirical processes were proved by Shorack and Wellner (cf. p. 470 in [42]) for $0 < p \leqq 2$ and by Csörgő and Horváth [10] for $0 < p < +\infty$ (see also Csörgő, Horváth and Shao [12]). In these papers strong approximation methods were used. Probability in Banach spaces techniques were adapted by Norvaiša in [35] and in [37] to prove this type of results in Lebesgue spaces and in Orlicz function spaces, respectively. In the proof of Theorem 2.1, as well as in those of all the other results of this paper, we use a combination of these methods, i.e. a combination of strong approximations with Probability in Banach spaces techniques. Due to counterexamples of Giné and Zinn [19], statement (3) of Theorem 2.1 cannot be derived from the central limit theorem for Banach space valued random variables without requiring in addition from a Banach space to satisfy Rosenthal's inequality (see Theorem 10.10 and Problem 15.8.7 in Ledoux and Talagrand [24] and Corollary 4.4 of Norvaiša [36]).

Let $\pi(X, Y)$ denote the Prohorov metric between distributions on a Banach function space F induced by stochastic processes X and Y with almost all sample paths in F. The following statement gives an estimate of the rate of convergence in (2.3).

PROPOSITION 2.2. *Let* F *be a separable Banach function space as in Theorem 2.1. Consider a uniform empirical process* E_n *and a Brownian bridge* B. *For each* $r > 2$ *there is a finite constant* $C(r)$ *such that, for all* $n \geqq 1$,

$$\pi(E_n, B) \leqq C(r)[1 \vee \|(I(1-I))^{1/r}\|]n^{-\frac{r/2-1}{1+r}}, \tag{2.4}$$

here and throughout the paper I *denotes the identical function* $I: t \to t$ *on* $(0, 1)$.

REMARK. (2.4) has meaning only if $\|(I(1-I))^{1/r}\| < +\infty$, i.e., if only the function

$$T \ni t \to (t(1-t))^{1/r} \in \mathsf{F}(T, m). \tag{2.5}$$

The estimate (2.4), for integers $r > 2$, under various additional assumptions also follows from corresponding results of Zolotarev [50] and of Bentkus and Račkauskas [4] for Banach space valued rv's. It is known also that the exponent of n in (2.4) is non-improvable in general (cf. Paulauskas and Račkauskas [39]).

It is well known that the empirical process E_n can be uniformly approximated also by a sequence of Poisson bridges $N_n = \{N_n(t);\, t \in (0,1)\}$, $n \geqq 1$, defined by

$$N_n(t) = n^{-1/2}[N(nt) - tN(n)], \quad t \in (0,1),$$

where N is a Poisson process with intensity parameter 1. Kac's modified empirical process defined by

$$K_n(t) = \sqrt{n}(G_n^*(t) - t), \quad t \in [0,1],$$

and

$$G_n^*(t) = \frac{1}{n}\sum_{i=1}^{\nu_n} 1_{(0,t]}(U_i), \quad t \in [0,1],$$

where $\{\nu_n;\, n \geqq 1\}$ is a sequence of Poisson real rv's with $E\nu_n = n$, independent of $\{U_i;\, i \geqq 1\}$, gives a convenient construction of such Poisson bridges N_n (cf. Section 7.0 in Csörgő and Révész [16]). Namely, for every $n \geqq 1$, we have

$$N_n(t) \stackrel{d}{=} K_n(t) - tK_n(1) = \frac{1}{\sqrt{n}}\sum_{i=1}^{\nu_n}(1_{(0,t]}(U_i) - t), \quad t \in (0,1). \tag{2.6}$$

In particular, according to this construction, N_n have almost all sample paths in a Banach function space simultaneously with the uniform empirical process E_n. Moreover, a simple calculation yields that the characteristic function of the induced distribution $\mathcal{L}(N_n)$ is equal to

$$\exp\left\{\int_{\mathbb{F}} (e^{i\langle x,f\rangle} - 1)n\mathcal{L}(Y(U)/\sqrt{n})(dx)\right\}, \qquad f \in \mathbb{F}^*,$$

where

$$Y(U)(t) = 1_{(0,t]}(U) - t, \qquad \forall\, t \in (0,1).$$

According to terminology from Probability in Banach spaces, $\mathcal{L}(N_n)$ is an accompanying Poisson law of $\mathcal{L}(E_n)$. Therefore, by Le Cam's theorem (cf., e.g., Theorems 3.4.8 and 3.4.9. in [2]), (2.3) is equivalent to

$$\mathcal{L}(N_n) \Rightarrow \mathcal{L}(B) \quad \text{on } \mathbb{F} \text{ weakly.}$$

Hence $\pi(E_n, N_n) \to 0$, as $n \to \infty$, whenever (2.3) holds true. The following statement says more about how close the distributions of E_n and N_n are to each other.

PROPOSITION 2.3. *Let F be a separable Banach function space as in Theorem 2.1. For any uniform empirical process E_n and for any Poisson bridge N_n we have:*

(1) *if $\sqrt{I(1-I)}\in\mathsf{F}$, i.e. if (2.1) holds, then there is a finite constant C such that, for all $n\geqq 1$,*

$$\pi(E_n, N_n) \leqq Cn^{-1/6};$$

(2) *if $(I(1-I))^{1/r}\in\mathsf{F}$ for some $r>2$, i.e. if (2.5) holds, then for all $q\in[2,r)$*

$$\pi(E_n, N_n) = o(n^{-\frac{q/4}{1+q}}).$$

We do not know how precise are these estimates. They are better than the corresponding result of Bakštys and Paulauskas [3] for Banach space valued rv's. Moreover, the arguments of the proof of Proposition 2.3 can be adapted to arbitrary Banach space valued rv's.

Turning now to invariance principles for partial sum processes, we note that the weak convergence of the weighted partial sum process in sup-norm was proved by O'Reilly [38] under the finite third moment condition. An extension of the Komlós, Major and Tusnády [23] and Major [29] approximation of partial sums to weighted sup-norm approximations which improve results of O'Reilly [38] in terms of the optimal class of weight functions, and in requiring the existence of $(2+\delta)$ moments was given by Csörgő and Horváth [9]. These results for the same class of weight functions and assuming two moments only were proved by Szyszkowicz [44] (for a summary along these lines we refer to [13]). To be more precise, for a given sequence of independent real rv's $X_1, X_2, \ldots$ with distribution function (df) F such that

$$\int x\,dF(x) = 0 \quad \text{and} \quad \int x^2 dF(x) = 1, \tag{2.7}$$

define the *partial sum process* $S_n = \{S_n(t);\, t\in(0,+\infty)\}$ by

$$S_n(t) = \frac{1}{\sqrt{n}}\sum_{i=1}^{[nt]} X_i, \quad t\in(0,+\infty).$$

Now we are ready to state an analogue of Theorem 2.1 for partial sum processes which is also an analogue of the L_p-norm approximation results of Szyszkowicz in [45] and in [46].

THEOREM 2.4. *Let $\mathsf{F} = (\mathsf{F}(T,m), \|\cdot\|)$, $T=(0,+\infty)$, be a separable Banach function space which does not contain l_∞^n's uniformly. The following statements about F are equivalent:*

(1) *the function*

$$T\ni t\to\sqrt{t}\in\mathsf{F}(T,m);$$

(2) *for every df F as in (2.7) one can construct a sequence of independent rv's with the df F and a sequence of Wiener processes $\{W_n;\ n \geqq 1\}$, in such a way that corresponding partial sum processes S_n, as well as Wiener processes W_n, have almost all sample paths in $\mathbb{F}$ and*

$$\|S_n - W_n\| = o_P(1); \tag{2.8}$$

(3) *there exists a Wiener distribution $\mathcal{L}(W)$ on $\mathbb{F}$ and for any sequence of independent rv's with the df F as in (2.7) the corresponding partial sum process S_n induces the distribution $\mathcal{L}(S_n)$ on $\mathbb{F}$ such that*

$$\mathcal{L}(S_n) \Rightarrow \mathcal{L}(W) \qquad \text{on } \mathbb{F} \text{ weakly.} \tag{2.9}$$

The following statement gives an estimate of the rate of convergence in (2.9).

PROPOSITION 2.5. *Let $\mathbb{F}$ be a separable Banach function space as in Theorem 2.4, and let*

$$v_q(t) := \begin{cases} t^{1/q}, & \text{if } 0 < t \leqq 1, \\ t^{1-1/q}, & \text{if } 1 \leqq t < +\infty. \end{cases}$$

Let F be a df as in (2.7) and such that

$$\int |x|^r dF(x) < \infty,$$

for some $r > 2$. Consider a partial sum process S_n corresponding to a sequence of independent rv's with the df F and a Wiener process W. There is a finite constant C such that, for all $n \geqq 1$,

$$\pi(S_n, W) \leqq C[1 \vee \|v_q\|]n^{-\frac{r(1/2-1/q)-1}{1+r}},$$

whenever

$$\frac{r}{q} \geqq 1 \qquad \text{or} \qquad r\left(\frac{1}{2} - \frac{1}{q}\right) < \frac{1}{1 - r/q}. \tag{2.10}$$

REMARK. From the proof it follows that assumptions (2.10) are superfluous if one considers Banach function spaces $\mathbb{F}(T, m)$ with $T = (0, t_0]$, $t_0 < +\infty$, instead of $T = (0, +\infty)$.

This is the first estimate of the rate of convergence of partial sums in the presence of weights. The exponent of n can be improved somewhat at the cost of much longer calculations. We will be content here with a weaker result because nothing is known about what its optimal version should be like.

We conclude with an invariance principle for a Poisson process. The proof of the implication (1) $\Rightarrow$ (3) of the following theorem, using tightness technique, was given by Norvaiša (cf. Theorem 3.6 in [34]).

THEOREM 2.6. *Let* $\mathsf{F} = (\mathsf{F}(T,m), \|\cdot\|)$, $T = (0,+\infty)$, *be a separable Banach function space which does not contain* l_∞^n*'s uniformly. The following statements about* F *are equivalent:*

(1) *the function*

$$T \ni t \to \sqrt{t} \in \mathsf{F}(T,m);$$

(2) *one can construct a net of Poisson processes* N_α *with intensity* α *and a net of Wiener processes* W_α *in such a way that centered Poisson processes* $N_\alpha - \alpha I$, *as well as Wiener processes* W_α, *have almost all sample paths in* F *and*

$$\|(N_\alpha - \alpha I)/\sqrt{\alpha} - W_\alpha\| = o_P(1), \quad \text{as} \quad \alpha \to +\infty; \tag{2.11}$$

(3) *there exists a Wiener distribution* $\mathcal{L}(W)$ *on* F *and a net of centered and scaled Poisson distributions with intensity* α *such that*

$$\mathcal{L}\left(\frac{N_\alpha - \alpha I}{\sqrt{\alpha}}\right) \Rightarrow \mathcal{L}(W) \quad \text{on } \mathsf{F} \text{ weakly, as} \quad \alpha \to +\infty.$$

To prove the implication (1) $\Rightarrow$ (2), we modify slightly a usual construction of the Poisson and Wiener processes with the desired joint distribution (see Proposition 5.5 below).

3. Banach function spaces

This section contains some facts about Banach function spaces to be used later on. This class of spaces includes among others Lebesgue spaces L_p, $1 \leqq p \leqq +\infty$, Orlicz, Lorentz, Marcinkiewicz and symmetric spaces.

Let (T,m) be a complete σ-finite separable measure space. Denote by $\mathsf{M} = \mathsf{M}(T,m)$ the linear space of all equivalence classes of m-measurable real-valued functions defined and finite m-a.e. on T. A map $\|\cdot\|: \mathsf{M} \to [0,+\infty]$ is called a *function norm* if

(1) $\|\cdot\|$ is a norm;

(2) $|f| \leqq |g|$ (m-a.e.) implies $\|f\| \leqq \|g\|$;

(3) if $A \subset T$ is of finite m-measure, then $\|1_A\| < +\infty$.

Given a function norm $\|\cdot\|$ on M, define the set

$$\mathsf{F}(T,m) := \{ f \in \mathsf{M}(T,m) : \|f\| < +\infty \}.$$

Then $\mathsf{F} = (\mathsf{F}(T,m), \|\cdot\|)$ is a normed linear space. If F is complete, it is called a *Banach function space*. We will assume further that all Banach function spaces are order complete (or Dedekind complete) and that m is a Radon σ-finite measure on a topological space T. The latter assumption in conjunction with the property (3) of the function norm $\|\cdot\|$ yield that

$\|1_K\| < +\infty$ for all compact subsets K of T. For notation not explained here we refer to Zaanen [49].

A Banach function space $\mathbb{F}$ is said to be *order continuous* (or to have an absolutely continuous norm) if for every $f \in \mathbb{F}$ and for every sequence $\{A_n;\ n \geqq 1\}$ of measurable subsets of T descending to a set of measure zero it follows that $\|f 1_{A_n}\| \to 0$, as $n \to \infty$. Note that L_p-spaces, with $1 \leqq p < +\infty$, are order continuous and that an Orlicz space L_ϕ is order continuous if and only if the Orlicz function ϕ satisfies Δ_2-condition.

Let $\mathbb{F}(T, m)$ be a Banach function space. The *associated space* (or Köthe dual) $\mathbb{F}' = \mathbb{F}'(T, m)$ is defined to be the set

$$\mathbb{F}' := \{\, g \in \mathbb{M}(T, m) \colon \int_T |fg|\, dm < +\infty \quad \forall f \in \mathbb{F} \,\}.$$

Each $g \in \mathbb{F}'$ defines a bounded linear functional on $\mathbb{F}$ by

$$f \longrightarrow \langle f, g\rangle := \int_T fg\, dm, \qquad f \in \mathbb{F}, \tag{3.1}$$

i.e. g may be considered as an element of a (topological) dual space $\mathbb{F}^*$.

Proposition 3.1. *Let $\mathbb{F} = (\mathbb{F}(T, m), \|\cdot\|)$ be a Banach function space with a measure space (T, m) as above. Then the following statements about $\mathbb{F}$ are equivalent:*

(1) *$\mathbb{F}$ is order continuous;*

(2) *$\mathbb{F}' = \mathbb{F}^*$ (isometrically) ;*

(3) *$\mathbb{F}$ is separable;*

(4) *$\mathbb{F}$ contains no subspace isomorphic with l_∞ (the Banach space of all bounded sequences with the sup-norm).*

Proof. (1) $\Leftrightarrow$ (2). See Theorem 1.2.3 in Luxemburg [28].

(1) $\Leftrightarrow$ (3). See Theorem 1.3.7 in Luxemburg [28].

(1) $\Leftrightarrow$ (4). See Lozanovskii [27].

To define Banach spaces which do not contain l_∞^n's uniformly, we proceed as follows. A Banach space $\mathbb{F}$ is said to *contain a subspace which is $(1+\epsilon)$-isomorphic to l_∞^n* if there exist $f_1, \ldots, f_n$ in $\mathbb{F}$ such that for all $\alpha = (\alpha_1, \ldots, \alpha_n) \in \Re^n$

$$\max_{1 \leqq i \leqq n} |\alpha_i| \leqq \Big\| \sum_{i=1}^n \alpha_i f_i \Big\| \leqq (1+\epsilon) \max_{1 \leqq i \leqq n} |\alpha_i|.$$

$\mathbb{F}$ contains l_∞^n's *uniformly*, or c_0 (the Banach space of all sequences converging to zero with the sup-norm) is finitely representable in $\mathbb{F}$, if it contains subspaces $(1+\epsilon)$-isomorphic to l_∞^n for all n and all $\epsilon > 0$. By Proposition 0.1 of Maurey and Pisier [33], in this definition it is sufficient to require the latter condition for some $\epsilon > 0$. Thus, if a Banach function space $\mathbb{F}$ does not contain l_∞^n's uniformly, then the statement (4) in Proposition 3.1 (and hence also all the others) holds true.

PROPOSITION 3.2. *Let* $\mathbb{F} = (\mathbb{F}(T,m), \|\cdot\|)$ *be a separable Banach function space. The following statements about* $\mathbb{F}$ *are equivalent:*

(1) $\mathbb{F}$ *does not contain* l_∞^n*'s uniformly;*

(2) *there exists* $p \in [1,+\infty)$ *such that* $\mathbb{F}$ *is p-concave, i.e. there exists a finite constant* $M_{(p)}$ *such that*

$$\left(\sum_{i=1}^{n} \|f_i\|^p\right)^{1/p} \leqq M_{(p)} \left\| \left(\sum_{i=1}^{n} |f_i|^p\right)^{1/p} \right\|,$$

for every choice of functions $f_1, \dots, f_n$ *in* $\mathbb{F}$*;*

(3) *There exist* $p \in [1,+\infty)$ *and a finite constant* $M_{(p)}$ *such that*

$$(E\|X\|^p)^{1/p} \leqq M_{(p)} \|(E|X|^p)^{1/p}\|,$$

for every measurable stochastic process $X = \{X(t); t \in T\}$ *with almost all sample paths in* $\mathbb{F}$.

REMARK 3.3. By Lemma 2.2 of James [21], it follows that if a Banach space $\mathbb{F}$ does not contain l_∞^n's uniformly, i.e. the statement (1) holds true, then $\mathbb{F}$ contains no subspace isomorphic with c_0.

PROOF. (1) $\Leftrightarrow$ (2). By Theorem of Maurey and Pisier [32], the Banach space $\mathbb{F}$ does not contain l_∞^n's uniformly if and only if there exists a finite number p such that $\mathbb{F}$ is of cotype p. By Proposition 1.f.3(i) and Corollary 1.f.9 from Lindenstrauss and Tzafriri [25], the latter fact is equivalent to the statement (2).

(2) $\Leftrightarrow$ (3). This is a consequence of Theorem 1.14 of Norvaiša [34]. Additional conditions for $\mathbb{F}$-valued rv's in that theorem are superfluous for measurable stochastic processes with sample paths in $\mathbb{F}$.

4. Distributions of Gaussian and Poisson processes

Let $X = \{X(t); t \in T\}$ be a measurable stochastic process on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$, and let $\mathbb{F} = (\mathbb{F}(T,m), \|\cdot\|)$ be a separable Banach function space. The question of X having almost all its sample paths in $\mathbb{F}$ is a part of a general problem of regularity of stochastic processes and, due to measurability of the function $\omega \to \|X(\cdot,\omega)\|$, it reduces to the question of having

$$\|X\| < +\infty \qquad \text{almost surely.} \tag{4.1}$$

This section contains a characterization of (4.1) for Gaussian and for centered Poisson processes when the Banach function space $\mathbb{F}$ does not contain l_∞^n's uniformly. Namely, it will be shown that (4.1) holds true for these processes if and only if

$$\| E|X| \| < +\infty, \tag{4.2}$$

where $E|X|(t) = E|X(t)|$, $t \in T$.

We start with the necessity of (4.2) and show it in two different ways. The first one is based on the following result of Csörgő, Horváth and Shao [12].

PROPOSITION 4.1. *Let $X = \{X(t); t \in T\}$ be a measurable stochastic process such that for some $r > 1$ there exists a finite constant $C = C(r)$ such that*

$$E|X(t)|^r \leqq C(E|X(t)|)^r, \qquad \forall t \in T, \tag{4.3}$$

and let f be an m-measurable function on T. Then

$$\int_T |X(t)f(t)|\, m(dt) < +\infty \qquad \textit{almost surely} \tag{4.4}$$

if and only if

$$\int_T E|X(t)||f(t)|\, m(dt) < +\infty. \tag{4.5}$$

REMARK. Condition (4.3) means that the family of random variables $\{X(t); t \in T\}$ satisfies the Marcinkiewicz–Paley–Zygmund condition. For some properties and examples of these families, we refer to Krakowiak and Szulga [22].

By Proposition 3.1, the duality between $\mathbb{F}$ and $\mathbb{F}^*$ is given by (3.1). Hence, (4.1) implies (4.4) for all $f \in \mathbb{F}^*$, and, by (4.5), it follows that $E|X| \in \in \mathbb{F}^{*\prime}$. Recall that a separable Banach function space $\mathbb{F}$ is perfect whenever $\mathbb{F}''(= \mathbb{F}^{*\prime}) = \mathbb{F}$. For example, all L_p-spaces with $1 \leqq p < +\infty$, are perfect Banach function spaces. Summarizing, we arrive at

COROLLARY 4.2. *Let $\mathbb{F}$ be a separable perfect Banach function space, and let $X = \{X(t); t \in T\}$ be a measurable stochastic process such that* (4.3) *holds true for some $r > 1$. Then* (4.1) *implies* (4.2).

These arguments in a converse direction do not imply the converse statement simply because the exceptional set in (4.4) depends on f in general. Nevertheless, Proposition 4.1 can be used to prove equivalence of (4.1) and (4.2) for L_p-spaces (cf. Csörgő, Horváth and Shao [12] for details). It is worthwile to mention also that, in the retrospect, the idea of the proof of Proposition 4.1 is similar to that used by Vahaniia [47] to characterize (4.1) for Gaussian sequences $X = \{X(n); n \geqq 1\}$ and $\mathbb{F} = l_p$, $1 \leqq p < +\infty$.

The advantage of Proposition 4.1 is that conditions on X are expressed directly via its distribution (cf. (4.3)). Other type of conditions used to prove (4.2) are based on certain integrability notions of X. More precisely, if (4.1) holds true, then a measurable stochastic process X defines, in a standard way, an F-valued rv $\dot{X}$. Note that, if a Banach function space F is perfect and does not contain subspace isomorphic with c_0, then (4.5) for all $f \in \mathsf{F}^*$ implies that $|\dot{X}|$ is Pettis integrable with the Pettis integral equal to $E|X|$. Thus, if we succeed in some way to show that $\dot{X}$ is Bochner integrable, i.e. if

$$E\|\dot{X}\| < +\infty, \tag{4.6}$$

then (4.1) implies (4.2). To follow this approach one needs to know the distribution of $\dot{X}$ given by the family of real rv's $\{\langle \dot{X}, f\rangle;\, f \in \mathsf{F}^*\}$. Since the duality between F and F^* is given by (3.1), the distribution of $\dot{X}$ is defined by distributions of the sample integrals

$$\int_T X(t) f(t)\, m(dt) = \langle \dot{X}, f\rangle =: X_f, \tag{4.7}$$

for all $f \in \mathsf{F}^*$.

Let $X = \{X(t); t \in T\}$ be a measurable Gaussian process with mean function μ, second order function k and covariance function r defined, respectively, by $\mu(t) = EX(t)$, $k(t) = EX^2(t)$ and $r(s,t) = EX(s)X(t) - EX(s)EX(t)$, for $s, t \in T$. Let us put also $\sigma^2(t) := r(t,t)$, for all $t \in T$. Clearly we have

$$\sigma \vee |\mu| \leqq \sqrt{k} \leqq \sigma + |\mu|, \qquad \text{on } T.$$

Moreover, due to the relation

$$E|X(t) - \mu(t)|^r = C(r)\sigma^r(t), \qquad \forall\, t \in T,$$

for any $r \in [1, +\infty)$, we have also

$$E|X| \leqq \sqrt{k} \leqq CE|X|, \qquad \text{on } T,$$

for some finite constant C. Thus, by Corollary 4.2, (4.1) implies (4.2) for the Gaussian process X whenever F is a perfect separable Banach function space.

Turning now to the second approach for proving the necessity of (4.2), we quote the following result.

PROPOSITION 4.3. *Let X be a measurable Gaussian process such that* (4.4) *holds for some m-measurable function f. Then the sample path integral X_f in* (4.7) *is a Gaussian rv.*

Different variants of the proof of this statement are due to Rajput [41], Liptser and Shiryaev (cf. p. 308 in [26]) and Vahaniia [48]. Thus, to identify the distribution of $\mathring{X}$, we only have to calculate EX_f^2 for all $f \in \mathbb{F}^*$. By Fubini's theorem, we have

$$EX_f^2 = \int_T \int_T f(t)f(s)r(t,s)m(dt)m(ds) = \langle Rf, f \rangle, \tag{4.8}$$

where

$$Rf(s) = \int_T f(t)r(t,s)\, m(dt), \qquad s \in T.$$

So, if the left-hand side of (4.8) is finite for all $f \in \mathbb{F}^*$, then R is a symmetric and positive operator from $\mathbb{F}^*$ into $\mathbb{F}^{*\prime}$. If in addition (4.1) holds true, then $R\colon \mathbb{F}^* \to \mathbb{F} \subset \mathbb{F}^{*\prime}$ (cf., e.g., p. 208 in [24]) and the characteristic function of the $\mathbb{F}$-valued rv $\mathring{X}$ is given by

$$Ee^{\imath\langle \mathring{X}, f\rangle} = \exp\{\imath\langle \mu, f\rangle - \langle Rf, f\rangle/2\}, \qquad f \in \mathbb{F}^*. \tag{4.9}$$

Due to the integrability of the norm of the Gaussian $\mathbb{F}$-rv $\mathring{X}$, (4.1) implies (4.2) whenever $\mathbb{F}$ is a separable Banach function space. The converse implication holds true whenever $\mathbb{F}$ does not contain l_∞^n's uniformly. More precisely, we have the following result of Gorgadze, Tarieladze and Čobanyan [20].

PROPOSITION 4.4. *Let $\mathbb{F}$ be a separable Banach function space. The following statements about $\mathbb{F}$ are equivalent*:

(1) *$\mathbb{F}$ does not contain l_∞^n's uniformly*;

(2) *every measurable Gaussian process $X \in \mathbb{F}$ almost surely if and only if $\sqrt{k} \in \mathbb{F}$*;

(3) *there exists an $\mathbb{F}$-valued rv $\mathring{X}$ with the characteristic function* (4.9) *if and only if $|\mu| \vee \sigma \in \mathbb{F}$*.

The proof of this statement is based on a theory of absolutely summing operators acting between Banach spaces. The implication (1) $\Rightarrow$ (2) also follows from Proposition 3.2.

Let $X_\alpha = \{X_\alpha(t);\, t \in T\}$, $T = (0, +\infty)$ be a measurable centered Poisson process with intensity parameter $\alpha > 0$, i.e. a measurable stochastic process with the characteristic function

$$E\exp\Big\{\imath \sum_{t\in T} a_t X_\alpha(t)\Big\}$$

$$= \exp\left\{ \alpha \int_T \big(e^{\imath \sum_{t\in T} a_t 1_{[s,\infty)}(t)} - 1 - \imath \sum_{t\in T} a_t 1_{[s,\infty)}(t)\big)ds \right\},$$

where $\{a_t;\, t \in T\}$ are real numbers all of which but finitely many are zero. To show the equivalence of (4.1) and (4.2) for X_α, one may argue like in case of a Gaussian process. For example, the property (4.3) of Proposition 4.1 follows from the following moment inequalities of (cf., e.g., Proposition 6.2 in [34]):

LEMMA 4.5. *Let η_u be a Poisson rv with intensity parameter $u > 0$ and let $1 \leqq p < +\infty$. Then there exist finite constants C_1, C_2, depending only on p, such that*

$$C_1 u e^{-u} \vee u^{p/2} \leqq E|\eta_u - u|^p \leqq C_2 u \vee u^{p/2}.$$

To prove that the induced $\mathbb{F}$-rv $\hat{X}_\alpha$ is Bochner integrable, one may argue as in Norvaiša [34] by showing that the characteristic function of $\hat{X}_\alpha$ is given by

$$E\exp\{\imath\langle \hat{X}_\alpha, f\rangle\} = \exp\left\{\int_{\mathbb{F}} \left[e^{\imath\langle x,f\rangle} - 1 - \imath\langle x, f\rangle\right] F_\alpha(dx)\right\}, \quad f \in \mathbb{F}^*, \tag{4.10}$$

where F_α is a Lévy measure of an infinitely divisible probability distribution on $\mathbb{F}$, defined on Borel subsets B by $F_\alpha(B) = \alpha\lambda(\kappa^{-1}(B))$, where $\kappa(t) = 1_{[t,\infty)}$ is a map from $(0, +\infty)$ into $\mathbb{F}$. Summarizing, by Theorems 2.7 and 2.9 of Norvaiša [34], we have:

PROPOSITION 4.6. *Let $\mathbb{F} = (\mathbb{F}(T, m), \|\cdot\|)$, $T = (0, +\infty)$ be a separable Banach function space which does not contain l^n_∞'s uniformly. Then the following statements hold true:*

(1) *every measurable centered Poisson process $X_\alpha \in \mathbb{F}$ a.s. if and only if $\sqrt{I} \wedge I \in \mathbb{F}$;*

(2) *there exists an infinitely divisible probability distribution on $\mathbb{F}$ with the characteristic function (4.10) if and only if $\sqrt{I} \wedge I \in \mathbb{F}$.*

5. Proofs

To prove the invariance principle for distributions induced by uniform empirical processes, we use the following part from Theorem 2.2 of M. Csörgő, S. Csörgő, Horváth, Mason [7]:

PROPOSITION 5.1. *There exists a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ with a sequence of uniform empirical processes $\{E_n;\ n \geqq 1\}$ and a sequence of Brownian bridges $\{B_n;\ n \geqq 1\}$, such that, for every $0 \leqq \nu < 1/4$,*

$$n^\nu \sup_{U_{n:1} \leqq t \leqq U_{n:n}} |E_n(t) - B_n(t)|/(t(1-t))^{1/2-\nu} = O_P(1). \tag{5.1}$$

PROOF OF THEOREM 2.1. (1) $\Rightarrow$ (2). We prove that the sequence of uniform empirical processes $\{E_n;\ n \geqq 1\}$, and the sequence of Brownian bridges $\{B_n;\ n \geqq 1\}$, from Proposition 5.1 satisfy statement (2) whenever (2.1) holds true. To see that the uniform empirical processes E_n have almost all sample paths in $\mathbb{F}$, we note that

$$E_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\psi(U_i, t) - a(t)), \qquad t \in (0,1),$$

where

$$\psi(u,t) = \begin{cases} -1_{[u,1/2)}, & \text{if } t \in (0,1/2), \\ 1_{[1/2,u)}, & \text{if } t \in [1/2,1), \end{cases}$$

for all $u \in (0,1)$, and

$$a(t) = \begin{cases} -t, & \text{if } t \in (0,1/2), \\ 1-t, & \text{if } t \in [1/2,1). \end{cases}$$

Hence, by the property (3) of the definition of the function norm $\|\cdot\|$, the uniform empirical process $E_n \in \mathbb{F}$ almost surely if and only if the function $I(1-I) \in \mathbb{F}$. Moreover, by Proposition 4.4, we have that any Brownian bridge process has almost all sample paths in $\mathbb{F}$ whenever $\sqrt{I(1-I)} \in \mathbb{F}$. Turning to the proof of the relation (2.2), we fix an arbitrary number $0 < \delta < 1/2$. Then we have, for all integers $n \geqq 1$,

$$\begin{aligned} \|E_n - B_n\| \leqq{}& \|(1_{(0,U_{n:1})} + 1_{(U_{n:n},1)})E_n\| + \|(1_{(0,U_{n:1})} + 1_{(U_{n:n},1)})B_n\| \\ &+ \|(1_{(0,\delta]} + 1_{[1-\delta,1)})\sqrt{I(1-I)}\| \sup_{U_{n:1} \leqq t \leqq U_{n:n}} |E_n(t) - B_n(t)|/\sqrt{t(1-t)} \\ &+ \|1_{(\delta,1-\delta)}\| \sup_{\delta \leqq t \leqq 1-\delta} |E_n(t) - B_n(t)| =: \sum_{i=1}^{4} I_i(n). \end{aligned} \tag{5.2}$$

Estimating $I_1(n)$, we have

$$I_1(n) \leqq \sqrt{nU_{n:1}}\|\sqrt{I}1_{(0,U_{n:1})}\| + \sqrt{n(1-U_{n:n})}\|\sqrt{1-I}1_{(U_{n:n},1)}\|.$$

Due to the order continuity of the Banach function space $\mathbb{F}$ (see Proposition 3.1), using the relations

$$nU_{n:1} = O_P(1) \qquad \text{and} \qquad n(1-U_{n:n}) = O_P(1),$$

it follows that

$$\lim_{n\to\infty} I_1(n) = 0 \quad \text{in probability.} \tag{5.3}$$

By Proposition 3.2, the Banach function space $\mathbb{F}$ is p-concave for some $1 \leqq p < +\infty$. Hence, for all $\epsilon > 0$ and all $z > 0$, we have

$$\begin{aligned} \mathbf{P}(\{ I_2(n) \geqq \epsilon \}) \leqq{}& \epsilon^{-p} M_{(p)}^p \|(1_{(0,z/n]} + 1_{[1-z/n,1)})\sqrt{I(1-I)}\|^p \\ &+ \mathbf{P}(\{ U_{n:1} > z/n \}) + \mathbf{P}(\{ U_{n:n} < 1 - z/n \}). \end{aligned}$$

By taking z sufficiently large, it follows that

$$\lim_{n\to\infty} I_2(n) = 0 \quad \text{in probability.} \tag{5.4}$$

Due to the order continuity of the Banach function space F once again, by (5.1) and taking δ sufficiently small, one can show that

$$\lim_{n\to\infty} (I_3(n) + I_4(n)) = 0 \quad \text{in probability.}$$

This, in conjunction with (5.2), (5.3) and (5.4), yields (2.2).

(2) $\Rightarrow$ (3). Since the distribution $\mathcal{L}(E_n)$ on F of the uniform empirical process E_n does not depend on a particular choice of a sequence of independent uniform rv's, statement (3) follows from that of (2) by Theorem 4.1 in [5].

(3) $\Rightarrow$ (4). Obvious.

(4) $\Rightarrow$ (1). By Remark 3.3, the Banach function space F does not contain a subspace isomorphic with c_0. Thus, by Proposition 5.1 of Pisier and Zinn [40] (cf. also Theorem 10.3 in [24]), there exists a Brownian bridge distribution $\mathcal{L}(B)$ on F. Now statement (1) follows from Proposition 4.4 and this completes also the proof of Theorem 2.1.

Next we show that Proposition 2.2 is a simple consequence of the construction given in the proof of Theorem 4.3.4 of Csörgő and Horváth [11]. This construction there was used to estimate a Prohorov metric between induced distributions on Skorohod space by weighted processes.

Proof of Proposition 2.2. By Strassen [43], it is sufficient to construct a sequence of uniform empirical processes $\{E_n;\, n \geqq 1\}$ and a sequence of Brownian bridges $\{B_n;\, n \geqq 1\}$ in such a way that E_n, as well as B_n, have almost all sample paths in F and

$$\mathbf{P}\left(\left\{\|E_n - B_n\| \geqq C(r)\|(I(1-I))^{1/r}\|n^{\frac{1-r/2}{1+r}}\right\}\right) \leqq C(r)n^{\frac{1-r/2}{1+r}} \tag{5.5}$$

for some finite constant $C(r)$ and all $n \geqq 1$. By using the construction of Komlós, Major and Tusnády [23] of the uniform empirical processes E_n and of the sequence of Brownian bridges B_n, it is shown in the proof of Theorem 4.3.4 of Csörgő and Horváth [11] that there exists a finite constant $C(r)$ such that

$$\mathbf{P}\left(\left\{\sup_{0<t<1} |E_n(t) - B_n(t)|/(t(1-t))^{1/r} \geqq C(r)n^{\frac{1-r/2}{1+r}}\right\}\right) \leqq C(r)n^{\frac{1-r/2}{1+r}}.$$

This establishes (5.5), and the proof of Proposition 2.2 is now complete.

To prove Proposition 2.3, we use a construction from the proof of Theorem 4.3.6 of Csörgő and Horváth [11] in conjunction with a necessary integrability condition for the central limit theorem to hold in an arbitrary Banach space.

PROOF OF PROPOSITION 2.3. Using Kac's construction (2.6) of Poisson bridges N_n, we have

$$E_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\nu_n} (1_{(0,t]}(U_i) - t) + R_n(t), \qquad t \in (0,1),$$

where

$$R_n(t) \stackrel{d}{=} \operatorname{sign}(n - \nu_n) \frac{1}{\sqrt{n}} \sum_{i=1}^{\eta_n} (1_{(0,t]}(U_i) - t), \qquad t \in (0,1),$$

and $\eta(n) = \eta_n = |\nu_n - n|$. By Theorem 2.1, condition $\sqrt{I(1-I)} \in \mathbb{F}$ implies (2.3). Hence, by Corollary 10.2 from [24],

$$\sup_{n \geqq 1} \sup_{z>0} z^2 \mathbf{P}(\{\|E_n\| > z\}) < +\infty.$$

Next, using Fubini's theorem, for all $x > 0$, we get

$$\begin{aligned} \mathbf{P}(\{\|E_n - N_n\| > x\}) &= \mathbf{P}(\{\|E_{\eta(n)}\| > x\sqrt{n/\eta_n}\}) \\ &\leqq n^{-1/2} x^{-2} \sup_{k \geqq 1} \sup_{z>0} z^2 \mathbf{P}(\{\|E_k\| > z\}) \sup_{k \geqq 1} E(\eta_k/\sqrt{k}). \end{aligned}$$

Choosing $x = n^{-1/6}$, by the Strassen [43] representation of Prohorov's metric we conclude statement (1) of Proposition 2.3. The proof of statement (2) goes along the lines of the proof of Theorem 4.3.6 of Csörgő and Horváth [11]. Only here, instead of the Birnbaum and Marshall inequality (see e.g. Inequality A.10.4 in [42]) for empirical process, one may use an extension of the Hájek and Rényi inequality for rv's with moments of order $r > 2$ (see e.g. Inequality A.10.3 in [42]). This, in turn, completes the proof of Proposition 2.3.

Similarly to the proof of the weighted L_p-norm results of Szyszkowicz (cf. [45] and [46]), the proof of our analogous invariance principle for distributions induced by the sequence of partial sum processes S_n uses the improvement of Strassen's invariance principle due to Major [30]:

PROPOSITION 5.2. *Let a df F as in (2.7) be given. Define*

$$\sigma_k^2 = \int_{-2^{n/2}}^{2^{n/2}} x^2 dF(x) - \left[\int_{-2^{n/2}}^{2^{n/2}} x \, dF(x) \right]^2 \quad \text{if} \quad 2^n \leqq k < 2^{n+1},\ n = 1, 2, \ldots.$$

A sequence of independent rv's $X_1, X_2, \ldots$ with the df F and a sequence of independent normal rv's $Y_1, Y_2, \ldots$ with $EY_k = 0$, $EY_k^2 = \sigma_k^2$ can be constructed in such a way that

$$\left| \sum_{i=1}^{n} (X_i - Y_i) \right| = o(n^{1/2}) \quad \text{almost surely.} \tag{5.6}$$

PROOF OF THEOREM 2.4. (1) $\Rightarrow$ (2). Let a df F as in (2.7) be given. Define a probability space $(\Omega, \mathcal{F}, P)$ to be a product of the probability space constructed by Proposition 5.2 and the probability space carrying a sequence of independent $N(0,1)$ rv's indexed by dyadic rational numbers. Then one can construct (cf. p. 22 in Csörgő and Révész [16]) a Wiener process W on $(\Omega, \mathcal{F}, \mathbf{P})$ in such a way that

$$W(n) = \sum_{i=1}^{n} Y_i/\sigma_i, \quad n \geqq 1.$$

We claim that the partial sum process S_n corresponding to the sequence X_1, $X_2, \ldots$, and the sequence of Wiener processes

$$W_n(t) := W(nt)/\sqrt{n}, \qquad n \geqq 1, \quad t \in (0, +\infty),$$

satisfy (2) whenever (1) holds true. To see that the partial sum processes S_n have almost all sample paths in F, we note that

$$(5.7) \quad |S_n(t)| \leqq \sqrt{t} \sup_{1 \leqq k < +\infty} k^{-1/2} \Big|\sum_{i=1}^{k}(X_i - Y_i)\Big| + |G_n(t)|, \qquad \forall\, t \in (0, +\infty),$$

where G_n is the partial sum process corresponding to a sequence $Y_1, Y_2, \ldots$. Hence G_n is a Gaussian process with the second order function

$$(5.8) \qquad EG_n^2(t) \leqq n^{-1} \sum_{i=1}^{[nt]} \sigma_i^2 \leqq t, \qquad \forall\, t \in (0, +\infty).$$

By Propositions 4.4, 5.2 and due to inequality (5.7), it follows that $S_n \in \mathsf{F}$ almost surely, for all $n \geqq 1$. Obviously, the same is true for the Wiener processes W_n, $n \geq 1$, by Proposition 4.4.

Turning now to proving (2.8), we fix real numbers $0 < \delta < D < +\infty$. Then, for all integers $n \geqq 1$, we have

$$\begin{aligned} I_n(\delta, D) &:= \sup_{\delta \leqq t \leqq D} |S_n(t) - W_n(t)| \\ &\leqq \sqrt{D} \sup_{n\delta \leqq k < +\infty} k^{-1/2} \left|\sum_{i=1}^{k}(X_i - Y_i)\right| + n^{-1/2} \max_{1 \leqq k \leqq Dn} \left|\sum_{i=1}^{k} Y_i(1 - 1/\sigma_i)\right| \\ &\quad + n^{-1/2} \sup_{0 \leqq t \leqq Dn} \sup_{0 \leqq s \leqq 1} |W(t+s) - W(t)|. \end{aligned}$$

Using Proposition 5.2, Lemma 1.2.1 of Csörgő and Révész [16] about the increments of a Wiener process and due to Kolmogorov's inequality, it follows that

$$(5.9) \qquad \lim_{n \to \infty} I_n(\delta, D) = 0 \qquad \text{in probability.}$$

Let us fix an arbitrary $\epsilon > 0$. Using Proposition 3.2 and the estimate (5.8), we get

$$\begin{aligned}
&\mathbf{P}(\{\|(S_n - W_n)(1_{(0,\delta)} + 1_{(D,+\infty)})\| > \epsilon\}) \\
&\leqq \mathbf{P}\left(\left\{\sup_{1\leqq k<+\infty} k^{-1/2}\left|\sum_{i=1}^{k}(X_i - Y_i)\right| \|\sqrt{I}(1_{(0,\delta)} + 1_{(D,+\infty)})\| > \epsilon/3\right\}\right) \\
&\quad + (\epsilon/3)^{-p}\|\,(E|G_n|^p)^{1/p}\,(1_{(0,\delta)} + 1_{(D,+\infty)})\|^p \\
&\quad + (\epsilon/3)^{-p}\|\,(E|W_n|^p)^{1/p}\,(1_{(0,\delta)} + 1_{(D,+\infty)})\|^p \\
&\leqq \mathbf{P}\left(\left\{\sup_{1\leqq k<+\infty} k^{-1/2}\left|\sum_{i=1}^{k}(X_i - Y_i)\right| > A\right\}\right) \\
&\quad + \mathbf{P}\left(\left\{A\|\sqrt{I}(1_{(0,\delta)} + 1_{(D,+\infty)})\| > \epsilon/3\right\}\right) \\
&\quad + 2(\epsilon/3)^{-p}\|\sqrt{I}(1_{(0,\delta)} + 1_{(D,+\infty)})\|^p.
\end{aligned}$$

Taking A, D to be sufficiently large, δ to be sufficiently small, one can make the right-hand side arbitrarily small. This, in conjunction with (5.9), yields (2.8).

(2) $\Rightarrow$ (3). Using a series representation of the partial sum process

$$S_n(t) = \sum_{i=1}^{\infty} \frac{X_i}{\sqrt{n}} 1_{[\frac{i}{n},+\infty]}(t), \qquad t \in (0,+\infty),$$

one may conclude that the distribution $\mathcal{L}(S_n)$ on $\mathbb{F}$ depends only on the df F. Therefore, (3) follows from (2) by Theorem 4.1 in [5].

(3) $\Rightarrow$ (1). Follows from Proposition 4.4, and this completes the proof of Theorem 2.4.

Similarly to Propositions 2.2 and 2.3, Proposition 2.5 will be shown to be a simple corollary of the next lemma, whose construction may also be used to estimate Prohorov's metric between distributions not necessarily sitting on Banach function spaces.

LEMMA 5.3. *Let F be a df and v_q be a function, both defined as in Proposition 2.5. One can construct a partial sum process S_n corresponding to a sequence of independent rv's with the df F and a sequence of Wiener processes W_n in such a way that there exists a finite constant C such that, for all $n \geqq 1$,*

$$\mathbf{P}\left(\left\{\sup_{0<t<\infty} |S_n(t) - W_n(t)|/v_q(t) \geqq Cz_n\right\}\right) \leqq Cz_n, \tag{5.10}$$

where $z_n = n^{-\frac{r(1/2-1/q)-1}{1+r}}$, whenever (2.10) *holds true.*

PROOF. By Komlós, Major and Tusnády [23], one can construct a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbf{P}})$, a sequence of independent rv's $X_1, X_2, \ldots$ with the df F and a sequence of standard normal rv's Y_1, Y_2, ... in such a way that, for all $n \geqq 1$,

$$\tilde{\mathbf{P}}\left(\left\{\max_{1\leqq k\leqq n}\left|\sum_{i=1}^{k}(X_i - Y_i)\right| > x_n\right\}\right) \leqq C_1 n x_n^{-r}, \tag{5.11}$$

provided $n^{1/r} \leqq x_n \leqq C_2\sqrt{n \log n}$, where C_1, C_2 are finite constants depending on the df F. Define a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ to be a product of the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbf{P}})$ and of the probability space carrying a sequence of independent $N(0, 1)$ rv's indexed by dyadic rational numbers. Then one can construct (cf. p. 22 in Csörgő and Révész [16]) a Wiener process $W = \{W(t);\ t \in (0, +\infty)\}$ on $(\Omega, \mathcal{F}, \mathbf{P})$ in such a way that

$$W(n) = \sum_{i=1}^{n} Y_i, \quad n \geqq 1.$$

We show that the partial sum process S_n corresponding to the sequence X_1, X_2, ..., and the sequence of Wiener processes

$$W_n(t) := W(nt)/\sqrt{n}, \qquad t \in (0, +\infty), \quad n \geqq 1,$$

satisfy (5.10). Indeed, by (5.11) and by Lemma 1.2.1 of Csörgő and Révész [16] concerning increments of a Wiener process, there is a finite constant C such that, for all $n \geqq 1$,

$$\begin{aligned}
&\mathbf{P}\left(\left\{\sup_{1/n\leqq t\leqq 1}|S_n(t) - W_n(t)|/v_q(t) \geqq 2z_n\right\}\right)\\
&\leqq \mathbf{P}\left(\left\{\max_{1\leqq k\leqq n}\left|\sum_{i=1}^{k}(X_i - Y_i)\right| \geqq n^{1/2-1/q} z_n\right\}\right)\\
&+ \mathbf{P}\left(\left\{\sup_{0<t\leqq n}\sup_{0<s<1}|W(t+s) - W(t)| \geqq n^{1/2-1/q} z_n\right\}\right)\\
&\leqq C_1 z_n^{-r} n^{1+r(1/q-1/2)} + n^{-2} \leqq C z_n.
\end{aligned} \tag{5.12}$$

Once again by (5.11) and by Lemma 1.2.1 of Csörgő and Révész [16], there is a finite constant C such that, for all $n \geqq 1$,

$$
\begin{aligned}
&\mathbf{P}\left(\left\{\sup_{1 \leqq t \leqq n^{r/q}} |S_n(t) - W_n(t)|/v_q(t) \geqq 2z_n\right\}\right) \\
&\leqq \mathbf{P}\left(\left\{\max_{1 \leqq k \leqq n^{1+r/q}} \left|\sum_{i=1}^{k}(X_i - Y_i)\right| \geqq n^{1/2}z_n\right\}\right) \\
&+ \mathbf{P}\left(\left\{\sup_{0<t<n^{1+r/q}} \sup_{0<s<1} |W(t+s) - W(t)| \geqq n^{1/2}z_n\right\}\right) \\
&\leqq z_n^{-r} n^{1+r(1/q-1/2)} + n^{-2} \leqq C z_n .
\end{aligned}
\tag{5.13}
$$

Using an extension of the Hájek–Rényi inequality to rv's with r moments (see e.g. Inequality A.10.3 in [42]), one can find a finite constant C such that, for all $n \geqq 1$,

$$
\begin{aligned}
&\mathbf{P}\left(\left\{\sup_{t \geqq n^{r/q}} |S_n(t)|/v_q(t) \geqq z_n\right\}\right) \\
&\leqq \mathbf{P}\left(\left\{\sup_{k \geqq n^{1+r/q}} \left|\sum_{i=1}^{k} X_i\right| /k^{1-1/q} \geqq n^{1/q-1/2}z_n\right\}\right) \\
&\leqq C \int |x|^r F(dx)\, z_n n^{r(1/q-1/2)(r/q-1)-1} .
\end{aligned}
\tag{5.14}
$$

To estimate similar probabilities for the Wiener process, it is sufficient to invoke the following inequality (cf., e.g., Lemma 4.2.1 in [11])

$$
\mathbf{P}\left(\left\{\sup_{0<t\leqq 1} |W(t)|/t^{1/q} \geqq z\right\}\right) \leqq C_1 \exp\{-C_2 z^2\},
\tag{5.15}
$$

for some finite constants C_1, C_2 and all $z \geqq 0$. Consequently, there exists a finite constant C such that, for all $n \geqq 1$,

$$
\begin{aligned}
&\mathbf{P}\left(\left\{\sup_{0<t\leqq 1/n} |W_n(t)|/v_q(t) \geqq C z_n\right\}\right) \\
&= \mathbf{P}\left(\left\{\sup_{0<t\leqq 1} |W(t)|/t^{1/q} \geqq C n^{1/2-1/q} z_n\right\}\right) \leqq n^{-2}.
\end{aligned}
\tag{5.16}
$$

Using (5.15) once again and scaling property of a Wiener process one can find another finite constant C such that, for all $n \geqq 1$,

$$\begin{aligned}(5.17)\qquad &\mathbf{P}\left(\left\{\sup_{t \geqq n^{r/q}} |W_n(t)|/v_q(t) \geqq C z_n\right\}\right) \\ &= \mathbf{P}\left(\left\{\sup_{0<t\leqq 1} |W(t)|/t^{1/q} \geqq C n^{r(1/2-1/q)/q} z_n\right\}\right) \leqq n^{-2}.\end{aligned}$$

Collecting now the estimates (5.12), (5.13), (5.14), (5.16), (5.17) and using assumption (2.8), we arrive at (5.10). This also completes the proof of Lemma 5.3.

Now we are ready for the

Proof of Proposition 2.5. Arguing as in the proof of Proposition 2.2, it is a simple consequence of Lemma 5.3 and the Strassen [43] representation of Prohorov's metric.

To prove our invariance principle for Poisson processes, we use the following strong approximation of partial sums due to Komlós, Major and Tusnády [23]:

Proposition 5.4. *Let F be a df for which*

$$\int x\, F(dx) = 0, \quad \int x^2 F(dx) = 1$$

and there is a $t_0 > 0$ such that

$$\int e^{tx} F(dx) < +\infty, \qquad \text{for} \quad |t| < t_0.$$

One can construct a sequence $X_1, X_2, \ldots$ of independent rv's with the df F and a sequence $Y_1, Y_2, \ldots$ of independent standard normal rv's sitting on a probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbf{P}})$ in such a way that, for all $z > 0$ and every integer $n \geqq 1$,

$$(5.18)\qquad \bar{\mathbf{P}}\left(\left\{\max_{1\leqq k\leqq n}\left|\sum_{i=1}^{k}(X_i - Y_i)\right| > C_1 \log n + z\right\}\right) < C_2 e^{-C_3 z},$$

where C_1, C_2, C_3 are positive constants, depending only on the df F, and C_3 can be taken as large as desired by choosing C_1 large enough.

An extension of this construction for suitable processes would incorporate a somewhat long and routine repetition of the arguments as stated in Csörgő [6, p. 21]. Usually this part of proof is omitted by saying that "without loss of generality a probability space is assumed to be so rich that all random variables and processes introduced later on can be defined on it". In our particular case this part of proof can be shortened by using the following standard measure-theoretical result in a different way.

LEMMA 5.5. *Let U_i, V_i, $i = 1, 2$, be Polish spaces, and let $\xi_i : U_i \to V_i$, $i = 1, 2$, be measurable maps. Let P_i be a probability measure on U_i, $i = 1, 2$, and let μ be a probability measure on the product space $V_1 \times V_2$ such that*

$$pr_{V_i}\mu = \mathcal{L}(\xi_i)\,(= \xi_i(P_i)), \quad i = 1, 2.$$

Then there exists a probability measure P on the product space $U_1 \times U_2$ such that

$$pr_{U_i}P = P_i, \quad i = 1, 2, \quad \text{and} \quad \xi_1 \otimes \xi_2(P) = \mu,$$

where $\xi_1 \otimes \xi_2((u_1, u_2)) = (\xi_1(u_1), \xi_2(u_2))$, i.e. μ is the joint distribution of (ξ_1, ξ_2) with respect to the probability P.

This is a special case of Theorem A.1 of de Acosta [1] (cf. also Lemma 4.4.4 in Csörgő and Révész [16]).

For a "traditional" proof of the following approximation result along the above mentioned lines, and for related results, we refer to [8], [14], [15] and [31].

PROPOSITION 5.6. *There exist Poisson and Wiener processes sitting on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ such that, for all positive real z and all real $D \geqq 2$, we have*

$$\mathbf{P}\left(\left\{\sup_{0<t\leqq D} |N(t) - t - W(t)| > C_1 \log D + z\right\}\right) < C_2 e^{-C_3 z}, \tag{5.19}$$

where C_1, C_2, C_3 are positive absolute constants and C_3 can be taken as large as desired by choosing C_1 large enough.

PROOF. Let $\pi(1)$ be a Poisson df with a parameter 1, and let $N(0,1)$ be a standard normal df. Since $\pi(1)$ has a moment generating function, by Proposition 5.4 one can construct a probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbf{P}})$ and two sequences $X_1, X_2, \ldots, Y_1, Y_2, \ldots$ of i.i.d. rv's with df's $\pi(1)$ and $N(0,1)$, respectively, such that for all $z > 0$ and every n (5.18) holds true. To construct Poisson and Wiener processes with the desired property we use Lemma 5.5 as follows. Take $V_1 = V_2 = \Re^\infty$, equipped with the product topology. Under this topology $\Re^\infty$ is a complete, separable, metrizable locally convex space. Moreover, $\Re^\infty$ is a Montel space and, hence, every sequence of rv's induces in a standard way an $\Re^\infty$-valued rv. Consider $S = \{\sum_{i=1}^k X_i;\ k \geqq 1\}$ and $T = \{\sum_{i=1}^k Y_i;\ k \geqq 1\}$ as rv's with values in $\Re^\infty$ and let μ denote a joint distribution of (S, T) on the product space $\Re^\infty \times \Re^\infty$. Take (U_i, P_i), $i = 1, 2$, to be a Polish space with Poisson $\bar{N}$ and Wiener $\bar{W}$ processes, respectively, on it and define rv's $\xi_i : U_i \to \Re^\infty$, $i = 1, 2$, by

$$\xi_1 := \{\bar{N}(k) - \bar{N}(k-1) - 1;\ k \geqq 1\}$$

and

$$\xi_2 := \{\bar{W}(k) - \bar{W}(k-1);\ k \geqq 1\},$$

respectively. Since $\bar{N}$ and $\bar{W}$ are processes with independent increments, the marginals of μ are the distributions of ξ_1 and ξ_2 on $\Re^\infty$, respectively. By Lemma 5.5, there exists a probability P on the product $\Omega := U_1 \times U_2$ such that the processes $N(t,\omega) := \bar{N}(t,u_1)$ and $W(t,\omega) := \bar{W}(t,u_2)$ are Poisson and Wiener processes, respectively, and, by Proposition 5.4, we have

$$\begin{aligned}
&\mathbf{P}\left(\left\{\max_{1\leqq k\leqq n}|N(k)-k-W(k)|>C_1\log n+z\right\}\right)\\
(5.20)\qquad &=\mu\left(\left\{(x,y)\in\Re^\infty\times\Re^\infty:\ \max_{1\leqq k\leqq n}|x_k-y_k|>C_1\log n+z\right\}\right)\\
&=\bar{\mathbf{P}}\left(\left\{\max_{1\leqq k\leqq n}\left|\sum_{i=1}^{k}(X_i-Y_i)\right|>C_1\log n+z\right\}\right)<C_2e^{-C_3z}.
\end{aligned}$$

Now, using the inequality

$$\begin{aligned}
\sup_{0<t\leqq D}|N(t)-t-W(t)| &\leqq \max_{1\leqq k\leqq [D]}|N(k)-k-W(k)|\\
&+\max_{1\leqq k\leqq [D]}|N(k+1)-N(k)|+1+\sup_{0<t\leqq D}\sup_{0\leqq s<1}|W(t+s)-W(t)|,
\end{aligned}$$

(5.19) follows from (5.20) and Lemma 1.2.1 of Csörgő and Révész [16] on the increments of a Wiener process. This completes the proof of Proposition 5.6.

PROOF OF THEOREM 2.6. (1) $\Rightarrow$ (2). Let N and W be the Poisson and the Wiener processes, respectively, constructed in Proposition 5.6. We claim that the nets

$$N_\alpha := N(\alpha\cdot),\quad \alpha>0,\qquad \text{and}\qquad W_\alpha := W(\alpha\cdot)/\sqrt{\alpha},\quad \alpha>0,$$

of the Poisson processes with intensity α and the Wiener processes, respectively, satisfy (2) whenever (1) holds true. By Propositions 4.4 and 4.6 we have that almost surely $N_\alpha-\alpha I\in\mathbb{F}$ and $W_\alpha\in\mathbb{F}$, for all $\alpha>0$. Turning now to proving (2.11), we fix real numbers $0<\delta<D<+\infty$. By Proposition 3.2, there exist $p\in[1,+\infty)$ and a finite constant $M_{(p)}$ such that

$$(5.21)\qquad \left(E\|W_\alpha\left(1_{(0,\delta]}+1_{[D,\infty)}\right)\|^p\right)^{1/p}\leqq M_{(p)}\|\sqrt{I}1_{[D,\infty)}\|.$$

By Lemma 4.5, there exists also a finite constant $C_{(p)}$ such that

$$(5.22)\qquad \left(E\|((N_\alpha-\alpha I)/\sqrt{\alpha})1_{[D,\infty)}\|^p\right)^{1/p}\leqq C_{(p)}M_{(p)}\|\sqrt{I}1_{[D,\infty)}\|.$$

By Proposition 5.6, we have

$$\begin{aligned}
&\sup_{\delta\leqq t\leqq D}|N_\alpha(t)-\alpha t-W(\alpha t)|/\sqrt{\alpha t}\\
(5.23)\qquad &=O\left(\frac{\log\alpha D}{\sqrt{\alpha\delta}}\right)=o(1),\qquad \text{almost surely, as } \alpha\to\infty,
\end{aligned}$$

and

$$\sup_{c/\alpha \leqq t \leqq \delta} |N_\alpha(t) - \alpha t - W(\alpha t)|/\sqrt{\alpha t} = O_P(1),$$

for every $c > 0$. Let Γ_1 be the time of the first jump of N_α. Then for each $\epsilon > 0$ there is a $c > 0$ such that

$$\mathbf{P}(\{\Gamma_1 < c\}) \leqq \epsilon. \tag{5.25}$$

Consequently, on the event $\{\Gamma_1 \geqq c\}$ we have

$$\begin{aligned} \|(N_\alpha - \alpha I)/\sqrt{\alpha} - W_\alpha\| \leqq \sqrt{c}\|\sqrt{I}1_{(0,c/\alpha]}\| \\ + \|\sqrt{I}1_{(0,\delta]}\| + \sup_{c/\alpha \leqq t \leqq \delta} |N_\alpha(t) - \alpha t - W(\alpha t)|/\sqrt{\alpha t} \\ + \|\sqrt{I}\| \sup_{\delta \leqq t \leqq D} |N_\alpha(t) - \alpha t - W(\alpha t)|/\sqrt{\alpha t}. \end{aligned} \tag{5.26}$$

Taking now δ sufficiently small and D sufficiently large, (2.11) follows from (5.21)–(5.26).

(2) $\Rightarrow$ (3). By Theorem 4.1 in [5].

(3) $\Rightarrow$ (1). By Proposition 4.4.

6. Appendix

Here we give a proof of the implication (2) $\Rightarrow$ (1) of Theorem C. Assume that (1) does not hold. Then by the the definition of the weight function w we have for all c

$$\lim_{n\to\infty} w(c/n)/\sqrt{n} = \infty \quad \text{or / and} \quad \lim_{n\to\infty} w(1-c/n)/\sqrt{n} = \infty. \tag{6.1}$$

We show only that the first relation in (6.1) contradicts the assumption (1). Proving that the second relation is impossible goes along the same lines and is therefore omitted. By taking $t = U_{n:1}/2$, it follows from statement (2) that

$$\sqrt{n}U_{n:1}\, w(U_{n:1}/2) = O_P(1). \tag{6.2}$$

Moreover, we have

$$1/(nU_{n:1}) = O_P(1). \tag{6.3}$$

Thus (6.2) and (6.3) yield

$$w(U_{n:1}/2)/\sqrt{n} = O_P(1). \tag{6.4}$$

By (6.3), one can find a positive constant C such that

$$\mathbf{P}(\{nU_{n:1} < C\}) \leqq 1/4, \tag{6.5}$$

for all sufficiently large n. By (6.4), one can find a finite constant A such that

$$\mathbf{P}(\{w(U_{n:1}/2)\sqrt{n} \geqq A\}) \leqq 1/4, \tag{6.6}$$

for all sufficiently large n. Hence, using the first relation in (6.1), (6.5) and (6.6), we arrive at the contradiction that

$$\begin{aligned} 1 &= \mathbf{P}(\{nU_{n:1} \geqq C\}) + \mathbf{P}(\{nU_{n:1} < C\}) \\ &\leqq \mathbf{P}(\{w(U_{n:1}/2) \geqq A\}) + \mathbf{P}(\{nU_{n:1} < C\}) \leqq 1/2, \end{aligned}$$

for all sufficiently large n. This completes the proof of the implication (2) $\Rightarrow$ $\Rightarrow$ (1) of Theorem C.

REFERENCES

[1] ACOSTA, A. DE, Invariance principles in probability for triangular arrays of B-valued random vectors and some applications, *Ann. Probab.* **10** (1982), 346–373. *MR* **83e**:60006

[2] ARAUJO, A. and GINÉ, E., *The central limit theorem for real and Banach valued random variables*, John Wiley & Sons, New York, 1980. *MR* **83e**:60003

[3] BAKŠTYS, G.A. and PAULAUSKAS, V.I., Convergence of distributions of sums of independent Banach-space-valued random elements with infinitely divisible laws, I, *Litovsk. Mat. Sb.* **26** (1986), 403–414 (in Russian). *MR* **88b**:60010. See also *Soviet Math. Dokl.* **35** (1987), 387–389.

[4] BENTKUS, V.YU. and RAČKAUSKAS, A., Estimates for the distance between sums of independent random elements in Banach spaces, *Teor. Veroyatnost. i Primenen.* **29** (1984), 46–64 (in Russian). *MR* **85j**:60006. See also *Theor. Probab. Appl.* **29** (1985), 50–65.

[5] BILLINGSLEY, P., *Convergence of probability measures*, Wiley, New York, 1968. *MR* **38** #1718

[6] CSÖRGŐ, M., *Quantile processes with statistical applications*, CBMS-NSF Regional Conference Series in Applied Mathematics, **42**, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1983. *MR* **86g**:60045

[7] CSÖRGŐ, M., CSÖRGŐ, S., HORVÁTH, L. and MASON, D.M., Weighted empirical and quantile processes, *Ann. Probab.* **14** (1986), 31–85. *MR* **87e**:60041

[8] CSÖRGŐ, M. and HORVÁTH, L., Asymptotic distributions of pontograms, *Math. Proc. Cambridge Philos. Soc.* **101** (1987), 131–139. *MR* **88m**:60057

[9] CSÖRGŐ, M. and HORVÁTH, L., Nonparametric methods for changepoint problems, *Handbook of Statistics*, Vol. 7, ed. by Krishnaiah P. R. and Rao C.R., North-Holland, Amsterdam, 1988, 403–425. *Zbl* **673**.62089

[10] CSÖRGŐ, M. and HORVÁTH, L., On the distributions of L_p norms of weighted uniform empirical and quantile processes, *Ann. Probab.* **16** (1988), 142–161. *MR* **89a**:60056

[11] CSÖRGŐ, M. and HORVÁTH, L., *Weighted approximations in probability and statistics*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, Wiley, Chichester, 1993. *MR* **94c**:60060

[12] CSÖRGŐ, M., HORVÁTH, L. and SHAO, Q.-M., Convergence of integrals of uniform empirical and quantile processes, *Stochastic Process. Appl.* **45** (1993), 283–294. *MR* **94f**:60052

[13] CSÖRGŐ, M., HORVÁTH, L., SHAO, Q.-M. and SZYSZKOWICZ, B., On the weighted asymptotics of partial sums and empirical processes of independent random variables, *Doeblin and modern probability* (Blaubeuren, 1991), Contemporary Math. **149**, Amer. Math. Soc., Providence, RI, 1993, 139–148. *MR* **94j**:60068

[14] Csörgő, M., Horváth, L. and Steinebach, J., Strong approximations for renewal processes, *C. R. Math. Rep. Acad. Sci. Canada* **8** (1986), 151–154. *MR* **87j**:60128

[15] Csörgő, M., Horváth, L. and Steinebach, J., Invariance principles for renewal processes, *Ann. Probab.* **15** (1987), 1441–1460. *MR* **88m**:60096

[16] Csörgő, M. and Révész, P., *Strong approximations in probability and statistics*, Probability and Mathematical Statistics, Akadémiai Kiadó, Budapest, 1981. *MR* **84d**:60050

[17] Csörgő, M., Shao, Q-M. and Szyszkowicz, B., A note on local and global functions of a Wiener process and some Rényi-type statistics, *Studia Sci. Math. Hungar.* **26** (1991), 239–259. *MR* **93i**:60078

[18] Dudley, R.M., An extended Wichura theorem, definitions of Donsker class, and weighted empirical distributions, *Probability in Banach spaces, V*, (Medford, Mass., 1984), Lecture Notes in Math., **1153**, Springer, Berlin–New York, 1985, 141–178. *MR* **87e**:60045

[19] Giné, E. and Zinn, J., Central limit theorems and weak laws of large numbers in certain Banach spaces, *Z. Wahrsch. Verw. Gebiete* **62** (1983), 323–354. *MR* **84j**:60017

[20] Gorgadze, Z.G., Tarieladze, V.I. and Čobanjan, S.A., Gaussian covariances in Banach sublattices of the space $L_0(T, \Sigma, \nu)$, *Dokl. Akad. Nauk SSSR* **241** (1978), 528–531. *MR* **80a**:60042. See also *Soviet Math. Dokl.* **19** (1978), 885–888.

[21] James, R.C., Uniformly non-square Banach spaces, *Ann. of Math.* (2) **80** (1964), 542–550. *MR* **30** #4139

[22] Krakowiak, W. and Szulga, J., Hypercontraction principle and random multilinear forms, *Probab. Theory Related Fields* **77** (1988), 325–342. *MR* **90e**:60007

[23] Komlós, J., Major, P. and Tusnády, G., An approximation of partial sums of independent RV's, and the sample DF. II, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **34** (1976), 33–58. *MR* **53** #6697

[24] Ledoux, M. and Talagrand, M., *Probability in Banach spaces. Isometry and processes*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Dritte Folge, **23**, Springer-Verlag, Berlin, 1991. *MR* **93c**:60001

[25] Lindenstrauss, J. and Tzafriri, L., *Classical Banach spaces II. Function spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Bd. 97, Springer-Verlag, Berlin–New York, 1979. *MR* **81c**:46001

[26] Liptser, R. Š. and Širjaev, A.N., *Statistics of random processes. I. General theory*, Applications of Mathematics, Vol. 5, Springer-Verlag, New York, 1977. *MR* **57** #14125; *II. Applications*, Applications of Mathematics, Vol. 6, Springer-Verlag, New York, 1978. *MR* **58** #7827

[27] Lozanovskiĭ, G.Ja., Isomorphic Banach lattices, *Sibirsk. Mat. Ž.* **10** (1969), 93–98. *MR* **39** #1942; See also *Siberian Math. J.* **10** (1969), 64–48.

[28] Luxemburg, W.A.J., *Banach function spaces*, Technische Hogeschool te Delft, 1955 (Thesis). *MR* **17**–285

[29] Major, P., The approximation of partial sums of independent RV's, *Z. Wahrscheinlichkeitstheor. Verw. Gebiete* **35** (1976), 213–220. *MR* **54** #3823

[30] Major, P., An improvement of Strassen's invariance principle, *Ann. Probab.* **7** (1979), 55–61. *MR* **82f**:60069

[31] Mason, D.M. and Zwet, W.R. van, A note on the strong approximation to the renewal process, *Publ. Inst. Statist. Univ. Paris* **32** (1987), 81–91. *MR* **88j**: 60066

[32] Maurey, B. and Pisier, G., Charactérisation d'une classe d'espaces de Banach par des propriétés de séries aléatoires vectorielles, *C.R.Acad. Sci. Paris Sér. A-B* **277** (1973), A687–A690. *MR* **48** #9352

[33] Maurey, B. and Pisier, G., Séries de variables aléatoires vectorielles indépendantes et propriétés géométriques des espaces de Banach, *Studia Math.* **58** (1976),

45–90. *MR* **56** #1388

[34] NORVAIŠA, R., Poisson and Brownian motion random functions with applications, Inst. of Math. and Inform., Preprint No. 13, Vilnius, 1991.

[35] NORVAIŠA, R., The central limit theorem for empirical and quantile processes in L_p-spaces, *Theor. Probab. Appl.* **37** (1992), 329–341.

[36] NORVAIŠA, R., The central limit theorem for L-statistics, Tech. rep. Series Lab. Res. Stat. Probab. **209**, Carleton Univ., Ottawa, 1992.

[37] NORVAIŠA, R., The central limit theorem for empirical and quantile processes in some Banach spaces, *Stochastic Process. Appl.* **46** (1993), 1–27. *MR* **94h**:60012

[38] O'REILLY, N., On the weak convergence of empirical processes in sup-norm metrics, *Ann. Probab.* **2** (1974), 642–651. *MR* **52** #4367

[39] PAULAUSKAS, V. and RAČKAUSKAS, A., *The accuracy of approximation in the central limit theorem in Banach spaces*, Mokslas, Vilnius, 1987 (in Russian). *MR* **89b**:60023. See also *Approximation theory in the central limit theorem. Exact results in Banach spaces*, Mathematics and its Applications (Soviet Series), **32**, Kluwer Academic Publishers, Dordrecht, 1989. *MR* **90h**:60006

[40] PISIER, G. and ZINN, J., On the limit theorems for random variables with values in the spaces L_p $(2 \leqq p < \infty)$, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **41** (1978), 289–304. *Zbl* **353**.60011 and **364**.60031

[41] RAJPUT, B.S., Gaussian measures on L_p spaces, $1 \leqq p < \infty$, *J. Multivariate Anal.* **2** (1972), 382–403. *MR* **49** #9896

[42] SHORACK, G.R. and WELLNER, J.A., *Empirical processes with applications to statistics*, Wiley Series in Probability and Mathematical Statistics, Probability and Mathematical Statistics, Wiley, New York, 1986. *MR* **88e**:60002

[43] STRASSEN, V., The existence of probability measures with given marginals, *Ann. Math. Statist.* **36** (1965), 423–439. *MR* **31** #1693

[44] SZYSZKOWICZ, B., Weighted L_p, $0 < p \leqq \infty$, convergence of partial sum type processes, Tech. Rep. Series Lab. Res. Stat. Probab. **191**, Carleton Univ., Ottawa, 1992.

[45] SZYSZKOWICZ, B., Weighted approximation of partial sum processes, Tech. Rep. Series Lab. Res. Stat. Probab. **206**, Carleton Univ., Ottawa, 1992.

[46] SZYSZKOWICZ, B., L_p-approximations of weighted partial sum processes, *Stochastic Process. Appl.* **45** (1993), 295–308. *MR* **94c**:60063

[47] VAHANIJA, N.N., Probabilistic distributions in linear spaces, *Sakharth. SSR Mecn. Akad. Gamothvl. Centr. Srom.* **10** (1971), no. 3, 155 pp. (in Russian). *MR* **45** #6047

[48] VAHANIJA, N.N., The correspondence between the concepts of a Gaussian measure and a Gaussian process, *Mat. Zametki* **26** (1979), 293–297 (in Russian). *MR* **81a**:60053

[49] ZAANEN, A.C., *Riesz spaces, II*, North-Holland Mathematical Library, **30**, Amsterdam, North-Holland Publishing Co., Amsterdam, 1983. *MR* **86b**:46001

[50] ZOLOTAREV, V.M., Ideal metrics in the problem of approximating the distributions of sums of independent random variables, *Teor. Verojatnost. i Primenen.* **22** (1977), 449–465 (in Russian). *MR* **56** #13306. See also *Theor. Probab. Appl.* **22** (1977), 433–449.

(Received August 1, 1994)

DEPARTMENT OF MATHEMATICS AND STATISTICS
CARLETON UNIVERSITY
1125 COLONEL BY DRIVE
OTTAWA, ONTARIO
K1S 5B6
CANADA
Fax: (613) 788-3822
gmurray@carleton.ca

INSTITUTE OF MATHEMATICS AND INFORMATICS
AKADEMIJOS 4
2600 VILNIUS
LITHUANIA
Fax: (3702) 359-209
r.norvaisa@mss.mii.lt

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 93–119

INTERMEDIATE ST. PETERSBURG SUMS

SÁNDOR CSÖRGŐ* and ROSSITZA DODUNEKOVA

Dedicated to Pál Révész for his sixtieth birthday

1. Introduction and results

Let X be Paul's gain in a St. Petersburg game, so that $\mathbf{P}\{X = 2^j\} = 2^{-j}$, $j \in \mathbf{N} := \{1, 2, \dots\}$. Then, setting $\lfloor x \rfloor = \max\{j \in \mathbf{Z} : j \leqq x\}$ and $\lceil x \rceil = \min\{j \in \mathbf{Z} : j \geqq x\}$, $x \in \mathbf{R}$, where $\mathbf{Z} := \{0, \pm 1, \pm 2, \dots\}$ and $\mathbf{R}$ is the set of real numbers, we have $F(x) := \mathbf{P}\{X \leqq x\} = 1 - 2^{-\lfloor \mathrm{Log}\, x \rfloor}$ for $x \geqq 2$, and for the corresponding left-continuous quantile function $Q(s) := \inf\{x : F(x) \geqq s\}$ one has $Q(s) = 2^{\lceil \mathrm{Log}(1/(1-s)) \rceil}$, $0 \leqq s < 1$. Here and throughout, Log stands for the logarithm to the base 2. Nicolaus Bernoulli's question in 1713, asking for Paul's 'fair price' for the game, led to the St. Petersburg paradox; for some historical references see [5] and [1].

Let $X_1, X_2, \dots$ be Paul's gains in a sequence of independent repetitions of the St. Petersburg game and, for each $n \in \mathbf{N}$, let $X_{1,n} \leqq \dots \leqq X_{n,n}$ be the order statistics pertaining to the sample $X_1, \dots, X_n$, so that $X_{1,n}, \dots, X_{n,n}$ are Paul's smallest, ..., largest gains in n St. Petersburg games. For the total gain $S_n := X_1 + \dots + X_n$ Feller proved in 1945 that $S_n/(n \,\mathrm{Log}\, n)$ converges in probability to 1; see [4]. One of the main difficulties with the St. Petersburg distribution is that it is not in the domain of attraction of any stable distribution. The first limiting distribution was found for the special subsequence $\{S_{2^n}\}$ by Martin-Löf [5]. Illustrating a 'probabilistic approach' to limit theorems surveyed in [3], the class of all possible subsequential limiting distributions for $\{S_n\}$ were found in [1] along with those for lightly, moderately and heavily trimmed variants of $\{S_n\}$, corresponding to certain portions of largest gains in n games renounced by Paul, and also those for moderately large sums of his largest winnings. Each of these are accompanied by a criteria for emerging along a given subsequence $\{n'\} \subset \mathbf{N}$. (These subsequences will be assumed unbounded below without further mention.) Meanwhile, the general theory has been completed in [2] with the description

1991 *Mathematics Subject Classification.* Primary 60F05.

Key words and phrases. St. Petersburg game, intermediate sums, asymptotic distributions.

* Partially supported by NSF Grant DMS-9208067.

of all possible subsequential limiting distributions for sums of intermediate order statistics arising out of the corresponding criteria to obtain them. The aim of the present paper is to accordingly round off the special St. Petersburg theory of limiting distributions by completely determining the class of all possible subsequential limiting laws for the sums of Paul's intermediately large winnings $X_{n-\lceil bk_n\rceil+1,n},\dots,X_{n-\lceil ak_n\rceil,n}$ in n games, i.e. for the sums $I_n(a,b):=I_n(a,b,k_n):=\sum_{i=\lceil ak_n\rceil+1}^{\lceil bk_n\rceil} X_{n+1-i,n}$, where $0<a<b$ are fixed numbers and $\{k_n\}$ is a sequence of positive numbers such that

$$k_n\to\infty \quad \text{and} \quad k_n/n\to 0 \quad \text{as} \quad n\to\infty, \tag{1.1}$$

with the corresponding criteria to obtain each of them along a given subsequence $\{n'\}\subset\mathbf{N}$. Even though the results are obtained as applications of some general theorems in [2], they do not come by easily. On the one hand, they show what is what underlies the general results of this type and what it takes to apply them to a non-trivial distribution that has been around for 282 years. On the other, they give an extra feel for and insight into the nature of the St. Petersburg game itself, particularly when seen in conjunction with all the other limit theorems in [1], informally described above.

The first goal is to determine the stochastic size of $I_n(a,b)$ and a workable centering. Let $-H(\cdot)$ be the left-continuous version of the right continuous function $Q(1-s)$, $0<s\leqq 1$, where $Q(\cdot)$ is the St. Petersburg quantile function as above, i.e. consider

$$H(s):=-\,2^{\lfloor \mathrm{Log}(1/s)\rfloor+1}, \qquad 0<s\leqq 1, \tag{1.2}$$

a left-continuous non-decreasing function, put

$$\mu_n(a,b):=\mu_n(a,b,k_n):=-\,n\int_{\lceil ak_n\rceil/n}^{\lceil bk_n\rceil/n} H(s)ds, \tag{1.3}$$

and let $\xrightarrow{\mathcal{D}}$ denote convergence in distribution.

THEOREM 1. *Let $0<a<b$, and suppose that (1.1) is satisfied for a sequence $\{k_{n'}\}$ of positive numbers along some subsequence $\{n'\}\subset\mathbf{N}$. If*

$$\frac{1}{A_{n'}}\left\{\sum_{i=\lceil ak_{n'}\rceil+1}^{\lceil bk_{n'}\rceil} X_{n'+1-i,n'}-C_{n'}\right\}\xrightarrow{\mathcal{D}} V^* \quad \text{as } n'\to\infty \tag{1.4}$$

for some constants $C_{n'}$ and $A_{n'}>0$ and a non-degenerate random variable V^*, then there exist a subsequence $\{n''\}\subset\{n'\}$ and a constant $0<\delta=\delta(\{n''\})<\infty$ such that

$$\frac{\sqrt{k_{n''}}\,2^{\lceil \mathrm{Log}(n''/k_{n''})\rceil}}{A_{n''}}\to\delta \quad \text{as } n''\to\infty \tag{1.5}$$

and

$$(1.6)\qquad \frac{1}{\sqrt{k_{n''}}\,2^{\lceil \mathrm{Log}(n''/k_{n''})\rceil}}\left\{\sum_{i=\lceil ak_{n''}\rceil+1}^{\lceil bk_{n''}\rceil} X_{n''+1-i,n''}-\mu_{n''}(a,b)\right\}\xrightarrow{\mathcal{D}} V$$

as $n''\to\infty$ *such that the distributional equality*

$$(1.7)\qquad V^*\stackrel{\mathcal{D}}{=}\delta V+\gamma$$

holds for some real constant γ.

To find all subsequential limiting types of distributions for $I_n(a,b)$, it is now possible to restrict attention to the asymptotic behavior only of the sequence

$$(1.8)\quad W_n:=\frac{1}{a_n}\left\{\sum_{i=\lceil ak_n\rceil+1}^{\lceil bk_n\rceil} X_{n+1-i,n}-\mu_n(a,b)\right\};\quad a_n:=\sqrt{k_n}\,2^{\lceil \mathrm{Log}(n/k_n)\rceil},$$

where for the centering sequence $\mu_n(a,b)$ in (1.3) elementary calculation from (1.2) gives

$$\mu_n(a,b)=n\left[\mathrm{Log}\,\frac{\lceil bk_n\rceil}{\lceil ak_n\rceil}+\beta\left(\frac{\lceil ak_n\rceil}{n}\right)-\beta\left(\frac{\lceil bk_n\rceil}{n}\right)\right]$$

for all n large enough, where, with $\langle s\rangle:=s-\lfloor s\rfloor$ standing for the fractional part of s, $\beta(s):=1+\langle \mathrm{Log}\, s\rangle-2^{\langle \mathrm{Log}\, s\rangle}$, $s>0$. The continuous function $\beta(s)$, $s>0$, has the property that $\beta\left(s\,2^k\right)=\beta(s)$ for every $s>0$ and $k\in\mathbf{Z}$, has range $[0,\,1-\{(1+\log\log 2)/\log 2\}]=[0,\,0.0860713\ldots]$, where log stands for the natural logarithm, takes 0 whenever $\mathrm{Log}\, s\in\mathbf{Z}$ and takes $1-\{(1+\log\log 2)/\log 2\}$ whenever $\langle \mathrm{Log}\, s\rangle=-(\log\log 2)/(\log 2)$.

For $x>0$ and $\{k_n\}$ as in (1.1), consider

$$(1.9)\qquad \begin{aligned} &0\leqq \gamma_x:=\lceil \mathrm{Log}\, x\rceil-\mathrm{Log}\, x<1,\\ &l_n:=\gamma_{n/k_n}=\lceil \mathrm{Log}(n/k_n)\rceil-\mathrm{Log}(n/k_n).\end{aligned}$$

Let $\{n'\}\subset\mathbf{N}$ be a subsequence and let $\{k_{n'}\}$ be a sequence of positive numbers such that (1.1) is satisfied along $\{n'\}$. Given $0<a<b<\infty$, for the separate choices of $c=a$ and $c=b$ we introduce four, mutually exclusive conditions:

c_1: $\lim_{n'\to\infty}\left(2^{\gamma_c-l_{n'}+1}-1\right)\sqrt{k_{n'}}=\infty$ and $\lim_{n'\to\infty}\left(2^{\gamma_c-l_{n'}}-1\right)\sqrt{k_{n'}}=-\infty$;

c_2: $\lim_{n'\to\infty}\left(2^{\gamma_c-l_{n'}+1}-1\right)\sqrt{k_{n'}}=\infty$ and $\lim_{n'\to\infty}\left(2^{\gamma_c-l_{n'}}-1\right)\sqrt{k_{n'}}=\infty$;

$c_3=c_3^u$: $\lim_{n'\to\infty}\left(2^{\gamma_c-l_{n'}}-1\right)\sqrt{k_{n'}}=u$ for some $-\infty<u<\infty$;

$c_4=c_4^v$: $\gamma_c=0$ and $\lim_{n'\to\infty}\left(2^{1-l_{n'}}-1\right)\sqrt{k_{n'}}=v$ for some $0\leqq v<\infty$, where the superscripts u and v for c_3 and c_4 will be used if the values of these parameters are specifically needed. We shall write $a_i \sqcap b_j$ when conditions a_i and b_j hold together, $1 \leqq i, j \leqq 4$. It turns out in Theorem 2 below that, in order for W_n to have non-degenerate subsequential limits in distribution, it is necessary that subsequential limits of l_n exist, along with some condition $a_i \sqcap b_j$ out of the possible twelve; $a_2 \sqcap b_4$, $a_3 \sqcap b_4$, $a_4 \sqcap b_2$ and $a_4 \sqcap b_3$ being impossible. For a given subsequence $\{n'\} \subset \mathbf{N}$, a subset of permissible conditions $a_i \sqcap b_j$ will then be delineated depending upon the values of

$$0 \leqq l := \lim_{n' \to \infty} l_{n'} \leqq 1 \quad \text{and} \quad \kappa := \kappa_{a,b} := \lceil \operatorname{Log} b \rceil - \lceil \operatorname{Log} a \rceil, \tag{1.10}$$

and their relationship with γ_a and γ_b in (1.9). Since $a < b$, the possible values for $\kappa = \gamma_b - \gamma_a + \operatorname{Log} b - \operatorname{Log} a$ are $0, 1, 2, \ldots$. If $\kappa = 0$, then necessarily $\gamma_a > \gamma_b$. We now list twenty-eight mutually exclusive conditions $\mathbf{C}_1, \ldots, \mathbf{C}_{28}$ expressing possible relationships among $\kappa, \gamma_a, \gamma_b$ and l, and each of these will be paired with a corresponding set $\mathcal{C}_1, \ldots, \mathcal{C}_{28}$ of permissible conditions $a_i \sqcap b_j$, $1 \leqq i, j \leqq 4$, some of which are the same.

$\mathbf{C}_1$: $\kappa \geqq 2$, $0 \leqq \gamma_a, \gamma_b < l \leqq 1$, but $l < 1$ if $\gamma_a \gamma_b = 0$;	$\mathcal{C}_1 = \{a_1 \sqcap b_1\}$,
$\mathbf{C}_2$: $\kappa \geqq 2$, $0 \leqq \gamma_a < l < \gamma_b$;	$\mathcal{C}_2 = \{a_1 \sqcap b_2\}$,
$\mathbf{C}_3$: $\kappa \geqq 2$, $0 \leqq \gamma_a < l < \gamma_b$;	$\mathcal{C}_3 = \{a_2 \sqcap b_1\}$,
$\mathbf{C}_4$: $\kappa \geqq 2$, $0 \leqq l < \gamma_a, \gamma_b$;	$\mathcal{C}_4 = \{a_2 \sqcap b_2\}$,
$\mathbf{C}_5$: $\kappa \geqq 2$, $0 = \gamma_a = \gamma_b = l$;	$\mathcal{C}_5 = \{a_1 \sqcap b_1, a_3 \sqcap b_3\}$,
$\mathbf{C}_6$: $\kappa \geqq 2$, $0 = \gamma_a = \gamma_b$, $l = 1$;	$\mathcal{C}_6 = \{a_1 \sqcap b_1, a_4 \sqcap b_4\}$,
$\mathbf{C}_7$: $\kappa \geqq 2$, $0 = \gamma_a = l < \gamma_b$;	$\mathcal{C}_7 = \{a_1 \sqcap b_2, a_3 \sqcap b_2\}$,
$\mathbf{C}_8$: $\kappa \geqq 2$, $0 = \gamma_a < \gamma_b = l$;	$\mathcal{C}_8 = \{a_1 \sqcap b_2, a_1 \sqcap b_3\}$,
$\mathbf{C}_9$: $\kappa \geqq 2$, $0 = \gamma_a < \gamma_b$, $l = 1$;	$\mathcal{C}_9 = \{a_1 \sqcap b_1, a_4 \sqcap b_1\}$,
$\mathbf{C}_{10}$: $\kappa \geqq 2$, $0 < \gamma_a = \gamma_b = l$;	$\mathcal{C}_{10} = \{a_1 \sqcap b_1, a_2 \sqcap b_2, a_3 \sqcap b_3\}$,
$\mathbf{C}_{11}$: $\kappa \geqq 2$, $0 < \gamma_a = l < \gamma_b$;	$\mathcal{C}_{11} = \{a_1 \sqcap b_2, a_2 \sqcap b_2, a_3 \sqcap b_2\}$,
$\mathbf{C}_{12}$: $\kappa \geqq 2$, $0 < \gamma_a < \gamma_b = l$;	$\mathcal{C}_{12} = \{a_1 \sqcap b_1, a_1 \sqcap b_2, a_1 \sqcap b_3\}$,
$\mathbf{C}_{13}$: $\kappa \geqq 2$, $0 = \gamma_b = l < \gamma_a$;	$\mathcal{C}_{13} = \{a_2 \sqcap b_1, a_2 \sqcap b_3\}$,
$\mathbf{C}_{14}$: $\kappa \geqq 2$, $0 = \gamma_b < \gamma_a = l$;	$\mathcal{C}_{14} = \{a_2 \sqcap b_1, a_3 \sqcap b_1\}$,
$\mathbf{C}_{15}$: $\kappa \geqq 2$, $0 = \gamma_b < \gamma_a$, $l = 1$;	$\mathcal{C}_{15} = \{a_1 \sqcap b_1, a_1 \sqcap b_4\}$,
$\mathbf{C}_{16}$: $\kappa \geqq 2$, $0 < \gamma_b = l < \gamma_a$;	$\mathcal{C}_{16} = \{a_2 \sqcap b_1, a_2 \sqcap b_2, a_2 \sqcap b_3\}$,
$\mathbf{C}_{17}$: $\kappa \geqq 2$, $0 < \gamma_b < \gamma_a = l$;	$\mathcal{C}_{17} = \{a_1 \sqcap b_1, a_2 \sqcap b_1, a_3 \sqcap b_1\}$,
$\mathbf{C}_{18}$: $\kappa = 1$, $\gamma_a > \gamma_b$;	$\mathcal{C}_{18} = \{a_1 \sqcap b_1, a_1 \sqcap b_4, a_2 \sqcap b_1, a_2 \sqcap b_2, a_2 \sqcap b_3, a_3 \sqcap b_1\}$,
$\mathbf{C}_{19}$: $\kappa = 1$, $\gamma_a = \gamma_b$;	$\mathcal{C}_{19} = \{a_1 \sqcap b_1, a_2 \sqcap b_2, a_3 \sqcap b_3, a_4 \sqcap b_4\}$,
$\mathbf{C}_{20}$: $\kappa = 1$, $l < \gamma_a < \gamma_b$;	$\mathcal{C}_{20} = \{a_2 \sqcap b_2\}$,
$\mathbf{C}_{21}$: $\kappa = 1$, $l = \gamma_a < \gamma_b$;	$\mathcal{C}_{21} = \{a_2 \sqcap b_2, a_3 \sqcap b_2\}$,
$\mathbf{C}_{22}$: $\kappa = 1$, $\gamma_a < \gamma_b = l$;	$\mathcal{C}_{22} = \{a_1 \sqcap b_1, a_1 \sqcap b_3\}$,

$\mathbf{C}_{23}$: $\kappa=1$, $0 \leqq \gamma_a < \gamma_b < l \leqq 1$, but $\gamma_a > 0$ if $l=1$; $C_{23}=\{a_1 \sqcap b_1\}$,
$\mathbf{C}_{24}$: $\kappa=1$, $0=\gamma_a<\gamma_b$, $l=1$; $C_{24}=\{a_1 \sqcap b_1, a_4 \sqcap b_1\}$,
$\mathbf{C}_{25}$: $\kappa=0$, $l=\gamma_b<\gamma_a$; $C_{25}=\{a_2 \sqcap b_1, a_2 \sqcap b_3\}$,
$\mathbf{C}_{26}$: $\kappa=0$, $\gamma_b<l<\gamma_a$; $C_{26}=\{a_2 \sqcap b_1\}$,
$\mathbf{C}_{27}$: $\kappa=0$, $\gamma_b<\gamma_a=l$; $C_{27}=\{a_2 \sqcap b_1, a_3 \sqcap b_1\}$,
$\mathbf{C}_{28}$: $\kappa=0$, $0=\gamma_b<\gamma_a$, $l=1$; $C_{28}=\{a_1 \sqcap b_4\}$.

Furthermore, for $c=a,b$, introduce

$$\lambda(c_i) := \begin{cases} -1, & \text{in cases } c_1,\ c_4, \text{ and } c_3=c_3^u \text{ with } u<0, \\ 0, & \text{in cases } c_2 \text{ and } c_3=c_3^u \text{ with } u \geqq 0, \end{cases} \tag{1.11}$$

for $x \in \mathbf{R}$, set $x^+ := \max(x,0)$ and $x^- := -\min(x,0)$ as usual, and for $x, u \in \mathbf{R}$, let

$$x^{(\operatorname{sgn} u)} := \begin{cases} x^+, & \text{if } u \geqq 0, \\ x^-, & \text{if } u<0. \end{cases}$$

Finally, let $\{W(t): t \geqq 0\}$ denote a standard Wiener process, and for $c=a,b$, define

$$V(c_1) := 0 =: V(c_2), \quad V(c_3^u) := \frac{[-W(c)-cu]^{(\operatorname{sgn} u)}}{2^{\lceil \operatorname{Log} c \rceil}}, \quad V(c_4^v) := \frac{[-W(c)-cv]^+}{2c}.$$

Recall the notation in (1.8), and let $\{k_{n'}\}$ be positive numbers for some $\{n'\} \subset \mathbf{N}$.

THEOREM 2. *Let $0<a<b$, and suppose that, for a sequence $\{k_{n'}\}$ of positive numbers, (1.1) is satisfied along a given subsequence $\{n'\} \subset \mathbf{N}$. If $l=\lim_{n' \to \infty} l_{n'}$ exists and for some $m=1,\dots,28$ the quadruplet $(\kappa_{a,b}, \gamma_a, \gamma_b, l)$ satisfies condition $\mathbf{C}_m$ and one of the conditions $a_i \sqcap b_j$ in the corresponding set C_m along $\{n'\}$, then*

$$\overline{W}_{n'} \xrightarrow{\mathcal{D}} V(l, a_i, b_j) := V(a_i) + \sum_{k=\lceil \operatorname{Log} a \rceil - \lambda(a_i)}^{\lceil \operatorname{Log} b \rceil - \lambda(b_j) - 1} \frac{W(2^{k-l})}{2^k} - V(b_j) \tag{1.12}$$

as $n' \to \infty$ and $V(l,a_i,b_j)$ is non-degenerate. Conversely, if $\overline{W}_{n'} \xrightarrow{\mathcal{D}} V$ for some non-degenerate random variable V as $n' \to \infty$, then every subsequence $\{n''\} \subset \{n'\}$ contains a further subsequence $\{n'''\} \subset \{n''\}$ such that $l''' = \lim_{n''' \to \infty} l_{n'''}$ exists and for some $m_{'''} = 1,\dots,28$ the quadruplet $(\kappa_{a,b}, \gamma_a, \gamma_b, l''')$ satisfies condition $\mathbf{C}_{m_{'''}}$ and one of the conditions $a_i''' \sqcap b_j'''$ in the corresponding set $C_{m_{'''}}$ along $\{n'''\}$, and hence $V \stackrel{\mathcal{D}}{=} V(l''', a_i''', b_j''')$. If $V \stackrel{\mathcal{D}}{=} V(l, a_i, b_j)$ for a permissible triplet (l, a_i, b_j) and the random variable $V(l,a_i,b_j)$ is not equal in distribution to $V(l^*, a_i^*, b_j^*)$ for any other permissible triplet (l^*, a_i^*, b_j^*), then the converse conclusion holds along the original subsequence $\{n'\}$.

The problem of sorting out the triplets (l, a_i, b_j) for which the condition of the last statement of the theorem may be violated appears to be difficult, but one has to count with such a lack of uniqueness in the parametric description of the possible limiting laws. It can be checked that there is no such lack of uniqueness in the subclass of limiting normal distributions, a member of which obtains whenever one of $\mathbf{C}_m$ is accompanied with a condition $a_i \sqcap b_j$, $i, j = 1, 2$, from the corresponding $\mathcal{C}_m$, $m = 1, \dots, 27$.

In general, all possible limits $V(a_i, b_j, l)$ arise. In fact, they can all be achieved along the single subsequence $\{n_r := 4^r\}_{r=1}^{\infty} \subset \mathbf{N}$, using $k_r := k_{n_r} := 2^{r+l_r(w)}$, where $l_r(w) := l^* + [w + \{(-1)^r/r\}]2^{-(\alpha r + l^*)/2}$, $r \in \mathbf{N}$, where $l^* \in [0, 1]$, $w \in \mathbf{R}$ and $\alpha \in \mathbf{R}$ are fixed parameters, with the convention that $w < 0$ when $l^* = 1$ and $\alpha > 0$. Then $0 < l_r < 1$ and $\operatorname{Log}(n_r/k_r) = \lceil \operatorname{Log}(n_r/k_r) \rceil - l_r(w)$ for all r large enough, and $l_r(w) \to l^*$ and $n_r/k_r = 2^{r - l_r(w)} \to \infty$ as $r \to \infty$. If now $0 \leqq \gamma < 1$ is fixed, then it is easy to see that $(2^{\gamma - l_r(w)} - 1)\sqrt{k_r} \to \xi(w)$ and if $l^* = 1$, then $(2^{1 - l_r(w)} - 1)\sqrt{k_r} \to \eta(w)$ as $r \to \infty$, where

$$\xi(w) = \begin{cases} -\infty, & \text{if } \gamma < l^*, \text{ or } \gamma = l^*,\ w > 0,\ \alpha < 1, \\ w \log 2, & \text{if } \gamma = l^*,\ \alpha = 1, \\ \infty, & \text{if } \gamma > l^*, \text{ or } \gamma = l^*,\ w < 0,\ \alpha < 1, \end{cases}$$

and

$$\eta(w) = \begin{cases} 0, & \text{if } \alpha > 1, \\ -w \log 2, & \text{if } \alpha = 1, \\ \infty, & \text{if } \alpha < 1. \end{cases}$$

Even though it is tedious to go through all the cases, it is now straightforward to see that, given any condition $\mathbf{C}_m$ and any condition $a_i \sqcap b_j$ in the corresponding set $\mathcal{C}_m$, it is possible to choose the parameters l^*, w and α so that, besides $\mathbf{C}_m$, the given $a_i \sqcap b_j$ obtains along $\{n_r\}$, $m = 1, \dots, 28$.

2. Proofs

First some preliminary considerations are presented. Throughout, $0 < a < b < \infty$ are fixed, and the fixed sequence $\{k_n\}$ of positive numbers satisfies (1.1). For $c = a, b$,

$$\operatorname{Log} \frac{n}{\lceil ck_n \rceil} = \operatorname{Log} \frac{n}{k_n} + \operatorname{Log} \frac{ck_n}{\lceil ck_n \rceil} - \operatorname{Log} c = \left\lceil \operatorname{Log} \frac{n}{k_n} \right\rceil - \lceil \operatorname{Log} c \rceil + \gamma_c - z_n(c),$$

where, with the notation in (1.9),

$$z_n(c) := l_n + \operatorname{Log} \frac{\lceil ck_n \rceil}{ck_n}, \qquad c = a, b. \tag{2.1}$$

Hence, substituting $s = \lceil ck_n \rceil / n$ into (1.2), we obtain

$$H\left(\frac{\lceil ck_n \rceil}{n}\right) = -2^{\left\lceil \operatorname{Log} \frac{n}{k_n} \right\rceil + 1} 2^{-\lceil \operatorname{Log} c \rceil + \lfloor \gamma_c - z_n(c) \rfloor},$$

so the increment of $H(\cdot)$ over the interval $(\lceil ak_n\rceil/n, \lceil bk_n\rceil/n)$ can be written as

$$\Delta_n(a,b) := H\left(\frac{\lceil bk_n\rceil}{n}\right) - H\left(\frac{\lceil ak_n\rceil}{n}\right) = 2^{\lceil \operatorname{Log}\frac{n}{k_n}\rceil + 1}\Delta_n^{\circ}(a,b), \tag{2.2}$$

where, since $H(\cdot)$ is non-decreasing,

$$\Delta_n^{\circ}(a,b) := 2^{-\lceil \operatorname{Log} a\rceil + \lfloor \gamma_a - z_n(a)\rfloor} - 2^{-\lceil \operatorname{Log} b\rceil + \lfloor \gamma_b - z_n(b)\rfloor} \geqq 0.$$

Setting $\Delta_n^*(a,b) := \Delta_n(a,b)$ if $\Delta_n(a,b) > 0$ and $\Delta_n^*(a,b) := 1$ if $\Delta_n(a,b) = 0$, for $c = a, b$ and all n large enough introduce the non-decreasing, left-continuous functions

$$\psi_n(c;x) := \begin{cases} \psi_n\left(c; -c\sqrt{k_n}/2\right), & \text{if } -\infty < x < -\frac{c\sqrt{k_n}}{2}, \\ \frac{1}{\Delta_n^*(a,b)}\left\{H\left(\frac{\lceil ck_n\rceil}{n} + x\frac{\sqrt{k_n}}{n}\right) - H\left(\frac{\lceil ck_n\rceil}{n}\right)\right\}, & \text{if } -\frac{c\sqrt{k_n}}{2} \leqq x \leqq \frac{c\sqrt{k_n}}{2}, \\ \psi_n\left(c; c\sqrt{k_n}/2\right), & \text{if } \frac{c\sqrt{k_n}}{2} < x < \infty, \end{cases}$$

for which $\psi_n(c;0) = 0$. Also, introduce the functions

$$\varphi_n(x) := \begin{cases} 0, & \text{if } -\infty < x < \frac{\lceil ak_n\rceil}{k_n}, \\ \frac{1}{\Delta_n^*(a,b)}\left\{H\left(\frac{xk_n}{n}\right) - H\left(\frac{\lceil ak_n\rceil}{n}\right)\right\}, & \text{if } \frac{\lceil ak_n\rceil}{k_n} \leqq x \leqq \frac{\lceil bk_n\rceil}{k_n}, \\ \varphi_n(\lceil bk_n\rceil/k_n), & \text{if } \frac{\lceil bk_n\rceil}{k_n} < x < \infty. \end{cases}$$

Clearly, $\varphi_n \equiv 0$ on $\mathbf{R}$ whenever $\Delta_n(a,b) = 0$, otherwise φ_n is a left-continuous distribution function on $\mathbf{R}$ for which $\varphi_n(x) = 1$ if $x \geqq \lceil bk_n\rceil/k_n$.

Since by (2.1) and (2.2),

$$\begin{aligned} \operatorname{Log}\frac{n}{\lceil ck_n\rceil + x\sqrt{k_n}} &= -\operatorname{Log}\left[\frac{\lceil ck_n\rceil}{n}\left(1 + \frac{x\sqrt{k_n}}{\lceil ck_n\rceil}\right)\right] = \\ &= -\operatorname{Log}\frac{\lceil ck_n\rceil}{ck_n} - \operatorname{Log} c - \operatorname{Log}\frac{k_n}{n} - \operatorname{Log}\left(1 + \frac{x\sqrt{k_n}}{\lceil ck_n\rceil}\right) = \\ &= -\left[l_n + \operatorname{Log}\frac{\lceil ck_n\rceil}{ck_n}\right] - \lceil \operatorname{Log} c\rceil + \gamma_c + \left\lceil \operatorname{Log}\frac{n}{k_n}\right\rceil - \\ &\quad - \operatorname{Log}\left(1 + \frac{x\sqrt{k_n}}{\lceil ck_n\rceil}\right) = \\ &= \left\lceil \operatorname{Log}\frac{n}{k_n}\right\rceil - \lceil \operatorname{Log} c\rceil + \gamma_c - z_n(c) - \operatorname{Log}\left(1 + \frac{x\sqrt{k_n}}{\lceil ck_n\rceil}\right), \end{aligned}$$

and so the increment $H(\{\lceil ck_n\rceil/n\} + \{x\sqrt{k_n}/n\}) - H(\lceil ck_n\rceil/n)$ is

$$2^{\lceil \operatorname{Log}\frac{n}{k_n}\rceil + 1 - \lceil \operatorname{Log} c\rceil}\left\{2^{\lfloor \gamma_c - z_n(c)\rfloor} - 2^{\lfloor \gamma_c - z_n(c) - \operatorname{Log}\left(1 + \{x\sqrt{k_n}/\lceil ck_n\rceil\}\right)\rfloor}\right\}$$

for all n large enough, for $c=a,b$ we obtain

$$\psi_n(c;x)=\frac{2}{2^{\lceil \operatorname{Log} c\rceil}}\frac{2^{\left\lceil \operatorname{Log}\frac{n}{k_n}\right\rceil}}{\Delta_n^*(a,b)}\psi_n^*(c;x), \tag{2.3}$$

where

$$\psi_n^*(c;x):=\begin{cases}\psi_n^*\left(c;-c\sqrt{k_n}/2\right), & \text{if } x<-\frac{c\sqrt{k_n}}{2},\\ 2^{\lfloor \gamma_c-z_n(c)\rfloor}-2^{\left\lfloor \gamma_c-z_n(c)-\operatorname{Log}\left(1+\left\{x\sqrt{k_n}/\lceil ck_n\rceil\right\}\right)\right\rfloor}, & \text{if } -\frac{c\sqrt{k_n}}{2}\leqq x\leqq\frac{c\sqrt{k_n}}{2},\\ \psi_n^*\left(c;c\sqrt{k_n}/2\right), & \text{if } \frac{c\sqrt{k_n}}{2}<x.\end{cases}$$

Also, since

$$\begin{aligned}\lfloor \operatorname{Log}(n/(xk_n))\rfloor &= -\lceil -\operatorname{Log}(n/(xk_n))\rceil\\ = -\lceil \operatorname{Log} x-\lceil \operatorname{Log}(n/k_n)\rceil+l_n\rceil &= \lceil \operatorname{Log}(n/k_n)\rceil-\lceil l_n+\operatorname{Log} x\rceil,\end{aligned}$$

we have

$$\varphi_n(x)=\frac{2^{\left\lceil \operatorname{Log}\frac{n}{k_n}\right\rceil+1}}{\Delta_n^*(a,b)}\varphi_n^*(x), \tag{2.4}$$

where

$$\varphi_n^*(x):=\begin{cases}0, & \text{if } x\leqq\frac{\lceil ak_n\rceil}{k_n},\\ 2^{-\lceil \operatorname{Log} a\rceil+\lfloor \gamma_a-z_n(a)\rfloor}-2^{-\lceil l_n+\operatorname{Log} x\rceil}, & \text{if } \frac{\lceil ak_n\rceil}{k_n}<x\leqq\frac{\lceil bk_n\rceil}{k_n},\\ 2^{-\lceil \operatorname{Log} a\rceil+\lfloor \gamma_a-z_n(a)\rfloor}-2^{-\lceil \operatorname{Log} b\rceil+\lfloor \gamma_b-z_n(b)\rfloor}, & \text{if } \frac{\lceil bk_n\rceil}{k_n}<x.\end{cases}$$

The equations (2.3) and (2.4) hold for all fixed $x\in\mathbf{R}$ if n is large enough. Let $\Longrightarrow$ denote weak convergence of functions, i.e. point-wise convergence of the corresponding functions at every continuity point of the limiting function in the interval to be indicated.

PROOF OF THEOREM 1. Suppose (1.4). By Theorem 2 in [2], there exist a subsequence $\{n''\}\subset\{n'\}$ and non-decreasing, left-continuous functions ψ_a and ψ_b on $\mathbf{R}$ such that $\psi_c(0)\leqq 0$, $\psi_c(0+)\geqq 0$ and for $c=a,b$.

$$\frac{\sqrt{k_{n''}}\Delta_{n''}^*(a,b)}{A_{n''}}\psi_{n''}(c;\cdot)\Longrightarrow\psi_c(\cdot)\ \text{ on }\ \mathbf{R}\quad\text{and}\quad\frac{\sqrt{k_{n''}}\Delta_{n''}(a,b)}{A_{n''}}\to\delta_1 \tag{2.5}$$

for some $0\leqq\delta_1<\infty$ as $n''\to\infty$. By the obvious inequalities

$$-2<\gamma_c-z_n(c)<1,\quad c=a,b, \tag{2.6}$$

the range of values of $\Delta_n^\circ(a,b)$ in (2.2) is a finite set that is bounded as a function of n. Thus one can choose a further subsequence of $\{n''\}$ along which

$\Delta_n^\circ(a,b)$ is a constant. Suppose that $\{n''\}$ is already such a subsequence of $\{n'\}$. We must consider two cases.

Case 1: $\Delta_{n''}^\circ(a,b) \equiv M > 0$. First we claim that δ_1 in (2.5) is necessarily positive. Suppose, to the contrary, that $\delta_1 = 0$. Then, using (2.2) and the notation in (1.8), we obtain $a_{n''}/A_{n''} \to 0$ as $n'' \to \infty$. Let x_c be any continuity point of $\psi_c(\cdot)$ in (2.5), $c = a, b$. Using (2.3), the convergence in (2.5) can be rewritten as

$$(2.7) \qquad \frac{a_{n''}}{A_{n''}} \psi_{n''}^*(c; x_c) \to 2^{\lceil \mathrm{Log}\, c \rceil - 1}\, \psi_c(x_c) \quad \text{as } n'' \to \infty.$$

Since, by (2.6), the sequence $\{\psi_{n''}^*(c; x_c)\}$ is bounded, it follows that $\psi_c(x_c) = 0$. Thus, $\psi_c \equiv 0$, $c = a, b$. Then, using the notation above (1.1) and in (1.3), by part (ii) of Theorem 1 in [2] it follows that $[I_{n''}(a,b) - \mu_{n''}(a,b)]/A_{n''} \xrightarrow{\mathcal{D}} 0$ as $n'' \to \infty$, which by the convergence of types theorem ([6], p. 42) implies that V^* in (1.4) is degenerate. Hence, indeed, $\delta_1 > 0$ in (2.5), and (1.5) follows from (2.2) and (2.5) with $\delta = \delta_1/(2M)$. Furthermore, (1.5) and (2.5) then imply that $\psi_{n''}(c; \cdot) \Longrightarrow \psi_c(\cdot)/\delta_1$ on $\mathbf{R}$, $c = a, b$, as $n'' \to \infty$. Since $\Delta_{n''}(a,b) = 2^{\lceil \mathrm{Log}(n''/k_{n''}) \rceil + 1} M$, this last convergence, an application of part (i) of Theorem 1 in [2], and a final application of the convergence of types theorem yield (1.6) and (1.7) along a further subsequence of the present $\{n''\}$.

Case 2: $\Delta_{n''}^\circ(a,b) \equiv 0$. In this case, $\delta_1 = 0$ in (2.5). Hence by part (ii) of Theorem 1 in [2], $[I_{n''}(a,b) - \mu_{n''}(a,b)]/A_{n''} \xrightarrow{\mathcal{D}} \overline{V}$, where $\overline{V}$ is degenerate if and only if $\psi_c \equiv 0$, $c = a, b$. However, by (1.4) and the convergence of types theorem $\overline{V}$ is non-degenerate, and so $\psi_c(x_c) \neq 0$ for some $x_c \in \mathbf{R}$ in (2.7) at least for one value of $c = a, b$. Since the sequence $\{\psi_{n''}^*(c; x_c)\}$ is bounded, we see that (1.5) must hold for some $\delta \in (0, \infty)$ along a further subsequence of the present $\{n''\}$, and the proof can, therefore, be completed as in Case 1. $\square$

The question of convergence of $\{W_n\}$ in (1.8), in Theorem 2, will be reduced to the question of weak convergence of the functions in (2.3) and (2.4). The latter problem will be solved in ten lemmas. To set the stage for the first of these, define, for $c = a, b$,

$$(2.8) \quad \psi_n^\circ(c; x) := \begin{cases} 2^{\lfloor \gamma_c - z_n(c) - \mathrm{Log}\left(1 + \{x\sqrt{k_n}/\lceil ck_n \rceil\}\right) \rfloor}, & \text{if } -\frac{c\sqrt{k_n}}{2} \leqq x \leqq \frac{c\sqrt{k_n}}{2}, \\ 0, & \text{if } |x| > \frac{c\sqrt{k_n}}{2}, \end{cases}$$

noting that $\psi_n^\circ(c; 0) - \psi_n^\circ(c; x) = \psi_n^*(c; x)$ for $x \in [-c\sqrt{k_n}/2, c\sqrt{k_n}/2]$. By (2.1),

$$\gamma_c - z_n(c) - \mathrm{Log}\left(1 + x\frac{\sqrt{k_n}}{\lceil ck_n \rceil}\right) = \gamma_c - l_n - \mathrm{Log}\, \frac{\lceil ck_n \rceil + x\sqrt{k_n}}{ck_n}.$$

Let $\varepsilon > 0$ be such that $\operatorname{Log}(1+\varepsilon) < 1$ and $\operatorname{Log}(1-\varepsilon) > \gamma_c - 1$, and for a fixed $x \in \mathbf{R}$ let n be so large that $|\{(\lceil ck_n \rceil + x\sqrt{k_n})/(ck_n)\} - 1| < \varepsilon$. Since $|\gamma_n - l_n| \leqq 1$, we have

$$-2 < -1 - \operatorname{Log}(1+\varepsilon) < \gamma_c - l_n - \operatorname{Log}\frac{\lceil ck_n \rceil + x\sqrt{k_n}}{ck_n} < \gamma_c - \operatorname{Log}(1-\varepsilon) < 1.$$

Hence the possible values of $\operatorname{Log}\psi_n^\circ(c;\cdot)$, for all n large enough, are only $-2,-1,0$. We have $\operatorname{Log}\psi_n^\circ(c;x) = -2$ if and only if $\gamma_c - l_n - \operatorname{Log}(\{\lceil ck_n \rceil + x\sqrt{k_n}\}/\{ck_n\}) < -1$ for all such n, or, equivalently,

$$x > [2^{\gamma_c - l_n + 1} - \{\lceil ck_n \rceil/(ck_n)\}]c\sqrt{k_n}.$$

Also, $\operatorname{Log}\psi_n^\circ(c;x) = -1$ if and only if $-1 \leqq \gamma_c - l_n - \operatorname{Log}(\{\lceil ck_n \rceil + x\sqrt{k_n}\}/\{ck_n\}) < 0$, which happens if and only if

$$[2^{\gamma_c - l_n} - \{\lceil ck_n \rceil/(ck_n)\}]c\sqrt{k_n} < x \leqq [2^{\gamma_c - l_n + 1} - \{\lceil ck_n \rceil/(ck_n)\}]c\sqrt{k_n}.$$

Finally, $\operatorname{Log}\psi_n^\circ(c;x) = 0$ if and only if $0 \leqq \gamma_c - l_n - \operatorname{Log}(\{\lceil ck_n \rceil + x\sqrt{k_n}\}/\{ck_n\})$, and this is the same as

$$x \leqq [2^{\gamma_c - l_n} - \{\lceil ck_n \rceil/(ck_n)\}]c\sqrt{k_n}.$$

Therefore,

$$\psi_n^\circ(c;x) := \begin{cases} 1, & \text{if } x \leqq u_n, \\ 1/2, & \text{if } u_n < x \leqq v_n, \\ 1/4, & \text{if } x > v_n, \end{cases} \tag{2.9}$$

where

$$u_n := u_n(c) := \left[2^{\gamma_c - l_n} - \frac{\lceil ck_n \rceil}{ck_n}\right]c\sqrt{k_n}$$

and

$$v_n := v_n(c) := \left[2^{\gamma_c - l_n + 1} - \frac{\lceil ck_n \rceil}{ck_n}\right]c\sqrt{k_n}.$$

Now we fix $c \in (0,\infty)$ arbitrarily; the choices $c = a$ and $c = b$ will be used later.

LEMMA 1. *Let $\{k_{n'}\}$ be positive numbers such that (1.1) is satisfied along $\{n'\} \subset \mathbf{N}$. Then the sequence $\{\psi_{n'}^\circ(c;\cdot)\}$ converges weakly on $\mathbf{R}$ if and only if one of the mutually exclusive conditions $c_1,\dots,c_4$ holds along $\{n'\}$. The corresponding limiting functions are:*

$$\psi_1^\circ(c;x) = 1/2, \quad x \in \mathbf{R}, \qquad \textit{if } c_1 \textit{ holds;}$$

$$\psi_2^\circ(c;x) = 1, \quad x \in \mathbf{R}, \qquad \textit{if } c_2 \textit{ holds;}$$

$$\psi_3^\circ(c;x) = \psi_{3,u}^\circ(c;x) = \begin{cases} 1, & \textit{for } x \leqq cu, \\ 1/2, & \textit{for } x > cu, \end{cases} \qquad \textit{if } c_3^u \textit{ holds;}$$

$$\psi_4^\circ(c;x) = \psi_{4,v}^\circ(c;x) = \begin{cases} 1/2, & \textit{for } x \leqq cv, \\ 1/4, & \textit{for } x > cv, \end{cases} \qquad \textit{if } c_4^v \textit{ holds.}$$

PROOF. Let $\overline{\mathbf{R}} := \{-\infty\} \cup \mathbf{R} \cup \{\infty\}$. First we show that the conditions

$$\lim_{n'\to\infty} u_{n'} = cu \quad \text{and} \quad \lim_{n'\to\infty} v_{n'} = cv \quad \text{for some } u, v \in \overline{\mathbf{R}}, \tag{2.10}$$

are necessary for the weak convergence of $\{\psi^{\circ}_{n'}(c;\cdot)\}$, where u_n and v_n are as in (2.9). Suppose, to the contrary, that at least one of them fails; the first, say. Then we see by (2.9) that $\psi^{\circ}_{n'}(c;x) = 1/2$ for infinitely many n' and $\psi^{\circ}_{n'}(c;x) = 1$ for infinitely many n' for any $x \in [\underline{u}, \overline{u}] \cap \mathbf{R}$, where $\underline{u} = \liminf_{n'\to\infty} u_{n'} < \limsup_{n'\to\infty} u_{n'} = \overline{u}$. Thus $\{\psi^{\circ}_{n'}(c;\cdot)\}$ cannot converge weakly, and the argument is the same if the second condition in (2.10) fails. So, indeed, (2.10) is necessary. But, since $[\{\lceil ck_n\rceil/(ck_n)\} - 1]\sqrt{k_n} = [\lceil ck_n\rceil - ck_n]/[c\sqrt{k_n}\,] \to 0$ as $n \to \infty$, the two conditions in (2.10) are equivalent to

$$\lim_{n'\to\infty} \left[2^{\gamma_c - l_{n'}} - 1\right]\sqrt{k_{n'}} = u \quad \text{and} \quad \lim_{n'\to\infty} \left[2^{\gamma_c - l_{n'}+1} - 1\right]\sqrt{k_{n'}} = v \tag{2.11}$$

for some $u, v \in \overline{\mathbf{R}}$. Thus the conditions in (2.11) are also necessary.

Next we show that they are sufficient as well. So, suppose that (2.11) holds for some $u, v \in \overline{\mathbf{R}}$. Then we have (2.10). Pick any $x \in \mathbf{R}$, $x \neq cu$, $x \neq cv$ if any of u and v is finite. Then, if $\varepsilon > 0$ is small enough, the interval $(x-\varepsilon, x+\varepsilon)$ contains neither $u_{n'}$ nor $v_{n'}$ for n' large enough, and hence we see by (2.9) that $\psi^{\circ}_{n'}(c;x)$ is constant for all n' large enough. So, (2.11) is in fact also sufficient for the weak convergence of $\{\psi^{\circ}_{n'}(c;\cdot)\}$. The whole sufficiency statement of the lemma is now trivial by straightforward inspection, including the stated forms of the limiting functions.

Returning to necessity, suppose that $\psi^{\circ}_{n'}(c;\cdot) \Longrightarrow \psi^{\circ}(c;\cdot)$ on $\mathbf{R}$, as $n' \to \infty$, for some function $\psi^{\circ}(c;\cdot)$, so that we have (2.11) for some $u, v \in \overline{\mathbf{R}}$. Let $\{n''\} \subset \{n'\}$ be any subsequence such that $0 \leqq l = \lim_{n''\to\infty} l_{n''\to\infty} \leqq 1$ exists. If $l < \gamma_c$, then $(u,v) = (\infty,\infty)$ in (2.11), so that c_2 holds along $\{n''\}$. If $l > \gamma_c$, then $u = -\infty$ in (2.11) and if, at the same time, $\gamma_c > 0$ or $\gamma_c = 0$ but $l < 1$, then $v = \infty$ in (2.11), so that c_1 holds along $\{n''\}$, whereas if $\gamma_c = 0$ and $l = 1$, then either $v = \infty$ in (2.11), so that c_1 holds along $\{n''\}$, or $0 \leqq v < \infty$ in (2.11), so that c_4^v holds along $\{n''\}$. If $l = \gamma_c$, then $v = \infty$ in (2.11) and c_1, c_2, or c_3^u takes place along $\{n''\}$, according as $u = -\infty$, $u = \infty$, or $u \in \mathbf{R}$.

The Bolzano–Weierstrass theorem and the above imply that if $\psi^{\circ}_{n'}(c;\cdot) \Longrightarrow \psi^{\circ}(c;\cdot)$ on $\mathbf{R}$, as $n' \to \infty$, then every subsequence of $\{n'\}$ contains a further subsequence along which exactly one of the conditions $c_1, \ldots, c_4$ is satisfied. If it is c_i for one of $i = 1, \ldots, 4$, then $\psi^{\circ}(c;\cdot) = \psi^{\circ}_i(c;\cdot)$ by the sufficiency part already established, and since the four limiting functions are different, we must have c_i along the whole $\{n'\}$. □

Notice that while both c_3 and c_4 yield the convergence of $\{l_{n'}\}$, c_1 and c_2 do not. That will come, in general, from the convergence of the φ_n functions in (2.4).

LEMMA 2. *Let $\{k_{n'}\}$ be positive numbers such that* (1.1) *is satisfied along $\{n'\} \subset \mathbf{N}$. Then the functions $\varphi^\circ_{n'}(x) := 2^{-\lceil l_{n'}+\operatorname{Log} x\rceil}$, $x>0$, converge weakly on $(0,\infty)$ as $n'\to\infty$ if and only if the limit $l=\lim_{n'\to\infty} l_{n'}$ exists. In this case, the left-continuous version of the limiting function is $\varphi^\circ_l(x) := 2^{-\lceil l+\operatorname{Log} x\rceil}$, $x>0$.*

PROOF. Suppose $\lim_{n'\to\infty} l_{n'} = l$ for some $l\in[0,1]$. If $D(\varphi)$ denotes the set of points of discontinuity of a function φ defined on a subset of $\mathbf{R}$, then $D(\varphi^\circ_l) = \{x>0 : l+\operatorname{Log} x\in\mathbf{Z}\} = \{x>0:\gamma_x - l\in\mathbf{Z}\}$. Clearly, $\lim_{n'\to\infty}\lceil l_{n'}+\operatorname{Log} x\rceil = \lceil l+\operatorname{Log} x\rceil$, and so $\varphi^\circ_{n'} \Longrightarrow \varphi^\circ_l$ on $(0,\infty)$ as $n'\to\infty$. Notice also that since $0\leqq\gamma_x<1$ as in (1.9), if $l\in[0,1)$, then $D(\varphi^\circ_l)=\{x>0:\gamma_x=l\}$ and $\varphi^\circ_1 = \varphi^\circ_0/2 \neq \varphi^\circ_l$. Hence $\varphi^\circ_{l_1}\neq\varphi^\circ_{l_2}$ for any $l_1,l_2\in[0,1]$ if $l_1\neq l_2$; the functions φ°_l, $0\leqq l\leqq 1$, are all different.

If $\varphi^\circ_{n'}\Longrightarrow\varphi$ for some function φ on $(0,\infty)$ as $n'\to\infty$, then any subsequence $\{n''\}\subset\{n'\}$ contains a further subsequence $\{n'''\}\subset\{n''\}$ such that $\lim_{n'''\to\infty} l_{n'''} = l$ for some $l\in[0,1]$, and hence, by sufficiency, $\varphi^\circ_{n'''}\Longrightarrow\varphi^\circ_l=\varphi$ on $(0,\infty)$ as $n'''\to\infty$. Thus l must be the same along all subsequences, and so $\lim_{n'\to\infty} l_{n'} = l$. □

Bringing $z_n(c) = l_n + \operatorname{Log}(\lceil ck_n\rceil/ck_n)$ in from (2.1), we need further specifications of conditions c^u_3 and c^v_4, $c=a,b$, of Theorem 2. Given a subsequence $\{n'\}\subset\mathbf{N}$, we shall write c^{u+}_3 if c^u_3 holds with $u>0$, c^{u-}_3 if c^u_3 holds with $u<0$, c^{0+}_3 if c^0_3 holds and $\lfloor\gamma_c - z_{n'}(c)\rfloor = 0$ for all n' large enough, c^{0-}_3 if c^0_3 holds and $\lfloor\gamma_c - z_{n'}(c)\rfloor=-1$ for all n' large enough, c^{v+}_4 if c^v_4 holds with $v>0$, c^{0+}_4 if c^0_4 holds and $\lfloor -z_{n'}(c)\rfloor = -1$ for all n' large enough, and c^{0-}_4 if c^0_4 holds and $\lfloor -z_{n'}(c)\rfloor=-2$ for all n' large enough. Recall also $\lambda(c_i)$, with the corresponding specifications of c_i, $i=1,\dots,4$, from (1.11).

LEMMA 3. *Let $\{k_{n'}\}$ be positive numbers such that* (1.1) *is satisfied along $\{n'\}\subset\mathbf{N}$. Suppose that $\lim_{n'\to\infty} l_{n'} = l$ for some $l\in[0,1]$ and that one of the conditions c_i, $i=1,\dots,4$, holds along $\{n'\}$.*

(1) *If c_1 holds, then $l\geqq\gamma_c$; if c^{u-}_3 holds, then $l=\gamma_c$; and if c^{v+}_4 holds, then $l=1$ and $\gamma_c=0$ necessarily. In all three cases $\lfloor\gamma_c - z_{n'}(c)\rfloor = \lambda(c_i) = -1$ for all n' large enough.*

(2) *If c_2 holds, then $l\leqq\gamma_c$; and if c^{u+}_3 holds, then $l=\gamma_c$. In both cases we have $\lfloor\gamma_c - z_{n'}(c)\rfloor = \lambda(c_i)=0$ for all n' large enough.*

(3) *If c^0_3 holds, then $l=\gamma_c$ and there exists a subsequence $\{n''\}\subset\{n'\}$ such that either c^{0+}_3 holds along $\{n''\}$ and $\lfloor\gamma_c - z_{n''}(c)\rfloor=\lambda(c^0_3)=0$ for all n''*

large enough, or c_3^{0-} holds along $\{n''\}$ and $\lfloor \gamma_c - z_{n''}(c) \rfloor = \lambda^*(c_3^{0-}) := -1$ for all n'' large enough.

(4) *If c_4^0 holds, then $l = 1$, $\gamma_c = 0$ necessarily, and there exists a subsequence $\{n''\} \subset \{n'\}$ such that either c_4^{0+} holds along $\{n''\}$ and $\lfloor \gamma_c - z_{n''}(c) \rfloor = = \lambda(c_4) = -1$ for all n'' large enough, or c_4^{0-} holds along $\{n''\}$ and $\lfloor -z_{n''}(c) \rfloor = \lambda^*(c_4^{0-}) := -2$ for all n'' large enough.*

Proof. The first statements in all (1)–(4) are trivial. To show the corresponding second statements, observe that

$$\lfloor \gamma_c - z_{n'}(c) \rfloor = \operatorname{Log} \psi_{n'}^{\circ}(c; 0-) = \operatorname{Log} \psi_{n'}^{\circ}(c; 0), \tag{2.12}$$

where $\psi_n^{\circ}(c;\cdot)$ is as in (2.8). Under c_i, by Lemma 1, $\psi_{n'}^{\circ}(c;\cdot) \Longrightarrow \psi_i^{\circ}(c;\cdot)$ on $\mathbf{R}$ as $n' \to \infty$, $i = 1, \dots, 4$. For the c_i in parts (1) and (2), $x = 0$ is a continuity point of $\psi_i^{\circ}(c;\cdot)$, and so $\psi_{n'}^{\circ}(c;0) \to \psi_i^{\circ}(c;0)$ as $n' \to \infty$. This, the definition of the limiting functions in Lemma 1 and (2.12) then yield the second statements. For (3), if c_3^0 holds, the limiting function $\psi_{3,0}^{\circ}(c;\cdot)$ jumps down from the value 1 at $x = 0$ to $1/2$, so that, by (2.12), $\lfloor \gamma_c - - z_{n'}(c) \rfloor = 0$ or -1 for all n' large enough, giving the desired alternative (both of which may be satisfied along different subsequences). Finally, if c_4^0 holds, the limiting function $\psi_{4,0}^{\circ}(c;\cdot)$ jumps down from the value $1/2$ to $1/4$, so that $\lfloor \gamma_c - z_{n'}(c) \rfloor = -1$ or -2 for all n' large enough, and the desired alternative follows again. □

To describe the limiting functions for ψ_n^* in (2.3), let $\psi_0^*(x) = 0$, $x \in \mathbf{R}$, and

$$\psi_{3,u}^{*+}(c;x) = \begin{cases} 0, & \text{if } x \le cu, \\ 1/2, & \text{if } x > cu, \end{cases} \qquad \psi_{3,u}^{*-}(c;x) = \begin{cases} -1/2, & \text{if } x \le cu, \\ 0, & \text{if } x > cu, \end{cases}$$

$$\psi_{4,v}^{*+}(c;x) = \begin{cases} 0, & \text{if } x \le cv, \\ 1/4, & \text{if } x > cv, \end{cases} \qquad \psi_{4,0}^{*-}(c;x) = \begin{cases} 1/4, & \text{if } x \le 0, \\ 0, & \text{if } x > 0, \end{cases}$$

where it is understood that $u \in [0,\infty)$ in $\psi_{3,u}^{*+}(c;\cdot)$, $u \in (-\infty, 0]$ in $\psi_{3,u}^{*-}(c;\cdot)$, and $v \in [0,\infty)$ in $\psi_{4,v}^{*+}(c;\cdot)$. All weak convergence statements in the next lemma are on $\mathbf{R}$.

Lemma 4. *Let $\{k_{n'}\}$ be positive numbers such that (1.1) is satisfied along $\{n'\} \subset \mathbf{N}$.*

(I) *Suppose that one of the conditions c_i, $i = 1, \dots, 4$, holds along $\{n'\}$.*

(I.1&2) *If c_1 or c_2 holds, then $\psi_{n'}^*(c;\cdot) \Longrightarrow \psi_0^*(\cdot)$; if c_3^{u+} holds, then $\psi_{n'}^*(c;\cdot) \Longrightarrow \psi_{3,u}^{*+}(c;\cdot)$; if c_3^{u-} holds, then $\psi_{n'}^*(c;\cdot) \Longrightarrow \psi_{3,u}^{*-}(c;\cdot)$; if c_4^{v+} holds, then $\psi_{n'}^*(c;\cdot) \Longrightarrow \psi_{4,v}^{*+}(c;\cdot)$ as $n' \to \infty$.*

(I.3) *If c_3^0 holds, then every subsequence $\{n''\} \subset \{n'\}$ contains a further subsequence $\{n'''\} \subset \{n''\}$ such that either c_3^{0+} holds along $\{n'''\}$ and*

$\psi^*_{n'''}(c;\cdot) \Longrightarrow \psi^{*+}_{3,0}(c;\cdot)$, *or* c_3^{0-} *holds along* $\{n'''\}$ *and* $\psi^*_{n'''}(c;\cdot) \Longrightarrow \psi^{*-}_{3,0}(c;\cdot)$ *as* $n''' \to \infty$.

(I.4) *If* c_4^0 *holds, then every subsequence* $\{n''\} \subset \{n'\}$ *contains a further subsequence* $\{n'''\} \subset \{n''\}$ *such that either* c_4^{0+} *holds along* $\{n'''\}$ *and* $\psi^*_{n'''}(c;\cdot) \Longrightarrow \psi^{*+}_{4,0}(c;\cdot)$, *or* c_4^{0-} *holds along* $\{n'''\}$ *and* $\psi^*_{n'''}(c;\cdot) \Longrightarrow \psi^{*-}_{4,0}(c;\cdot)$ *as* $n''' \to \infty$.

Furthermore, for every $u \in \mathbf{R}$ *and* $v \in [0,\infty)$,

$$\int_0^{-W(c)} \psi^{*+}_{3,u}(c;x)\,dx = \frac{[-W(c)-cu]^+}{2} = \frac{[-W(c)-cu]^{(\operatorname{sgn} u)}}{2},$$

$$\int_0^{-W(c)} \psi^{*-}_{3,u}(c;x)\,dx = \frac{[-W(c)-cu]^-}{2} = \frac{[-W(c)-cu]^{(\operatorname{sgn} u)}}{2},$$

$$\int_0^{-W(c)} \psi^{*+}_{4,v}(c;x)\,dx = \frac{[-W(c)-cv]^+}{4},$$

$$\int_0^{-W(c)} \psi^{*-}_{4,0}(c;x)\,dx = \frac{[-W(c)]^-}{4},$$

where $\{W(t): t \geqq 0\}$ *is a standard Wiener process as in* (1.12).

(II) *Suppose that* $\psi^*_{n'}(c;\cdot) \Longrightarrow \psi^*(\cdot)$ *for some function* $\psi^*(\cdot)$ *on* $\mathbf{R}$ *as* $n' \to \infty$. *If* $\psi^*(\cdot) \not\equiv 0$, *then one of the seven conditions* c_3^{u-}, c_3^{u+}, c_3^{0-}, c_3^{0+}, c_4^{v+}, c_4^{0-} *and* c_4^{0+} *listed before Lemma 3 are satisfied along the same* $\{n'\}$, *and we have*

$$\frac{2}{2^{\lceil \operatorname{Log} c\rceil}} \int_0^{-W(c)} \psi^*(x)\,dx = \begin{cases} V(c_i), & \text{for } c_3^{u-}, c_3^{u+}, c_3^{0+}, c_4^{v+}, c_4^{0+}, \\ V(c_3) + \frac{W(c)}{2^{\lceil \operatorname{Log} c\rceil}}, & \text{for } c_3^{0-}, \\ V(c_4) + \frac{W(c)}{2^{\lceil \operatorname{Log} c\rceil+1}}, & \text{for } c_4^{0-}, \end{cases}$$

where the random variables $V(c_i)$, $i=1,\dots,4$, *are as in Theorem 2. If, on the other hand,* $\psi^*(\cdot) \equiv 0$, *then every subsequence of* $\{n'\}$ *contains a further subsequence along which either* c_1 *or* c_2 *holds, and* $2^{1-\lceil \operatorname{Log} c\rceil} \int_0^{-W(c)} \psi^*(x)\,dx =$ $= 0 = V(c_1) = V(c_2)$.

Proof. Recall that $\psi^*_{n'}(c;x) = \psi^\circ_{n'}(c;0) - \psi^\circ_{n'}(c;x)$, $x \in [-c\sqrt{k_{n'}}/2, c\sqrt{k_{n'}}/2]$, as noted at (2.8). For c_1 and c_2 the statements in (I.1&2) follow directly from Lemma 1. As pointed out before Lemma 2, both c_3 and

c_4 imply that $l = \lim_{n'} l_{n'}$ exists, and so the other three statements in (I.1&2) follow by combining Lemma 1 and the corresponding statements in Lemma 3(1&2). Similarly, Lemma 1 combined with Lemma 3(3) and Lemma 3(4) results in the statements in (I.3) and (I.4), respectively.

To prove (II), recall also that $\lfloor \gamma_c - z_{n'}(c) \rfloor = \operatorname{Log} \psi^\circ_{n'}(c;0) \in \{-2,-1,0\}$ for all n' large enough. Suppose first that $\lfloor \gamma_c - z_{n'}(c) \rfloor = m_{\prime}$ for all n' large enough and some fixed $m_{\prime} \in \{-2,-1,0\}$. Then $\psi^\circ_{n'}(c;\cdot) = 2^{m_{\prime}} - \psi^*_{n'}(c;\cdot) \Longrightarrow$ $\Rightarrow 2^{m_{\prime}} - \psi^*(c;\cdot)$ as $n' \to \infty$. Then, by Lemma 1, one of $c_1, \dots, c_4$ holds along $\{n'\}$. If $\psi^*(\cdot) \not\equiv 0$, then by the already established (I.1–4) this condition is one of $c_3^{u-}, c_3^{u+}, c_3^{0-}, c_3^{0+}, c_4^{u+}, c_4^{0-}$ and c_4^{0+}, whereas if $\psi^*(\cdot) \equiv 0$, it has to be one of c_1 and c_2.

Next, suppose that for two subsequences $\{n''\} \subset \{n'\}$ and $\{n'''\} \subset \{n'\}$ we have $\lfloor \gamma_c - z_{n''}(c) \rfloor = m_{\prime\prime}$ for all large n'' and $\lfloor \gamma_c - z_{n'''}(c) \rfloor = m_{\prime\prime\prime}$ for all large n''', so that $\psi^\circ_{n''}(c;\cdot) \Longrightarrow 2^{m_{\prime\prime}} - \psi^*(c;\cdot) =: \psi^\circ_{\prime\prime}(c;\cdot)$ and $\psi^\circ_{n'''}(c;\cdot) \Longrightarrow 2^{m_{\prime\prime\prime}} -$ $- \psi^*(c;\cdot) =: \psi^\circ_{\prime\prime\prime}(c;\cdot)$ on $\mathbf{R}$, as $n'' \to \infty$ and $n''' \to \infty$, respectively, for some different $m_{\prime\prime}, m_{\prime\prime\prime} \in \{-2,-1,0\}$. As above, a condition $c_{i''}$ must be satisfied along $\{n''\}$, and a condition $c_{i'''}$ must be satisfied along $\{n'''\}$ for some $i'', i''' =$ $= 1, \dots, 4$. Since $\psi^\circ_{\prime\prime}(c;x) - \psi^\circ_{\prime\prime\prime}(c;x) = 2^{m_{\prime\prime}} - 2^{m_{\prime\prime\prime}}$, a constant for all $x \in \mathbf{R}$, it follows from the description of the possible limiting functions $\psi^\circ_i(c;\cdot)$, $i =$ $= 1, \dots, 4$, in Lemma 1 that both $\psi^\circ_{\prime\prime}(c;\cdot)$ and $\psi^\circ_{\prime\prime\prime}(c;\cdot)$ have to be constant on the whole $\mathbf{R}$. Hence, necessarily, both i'' and i''' is one of 1 and 2. (And $\psi^*(\cdot) \equiv 0$ in both cases.) Thus, both statements in (II) are established. All the integral statements in (I) follow by elementary calculations, and those in (II) come from the ones in (I). □

Now we aim a closer look at the functions φ_n in (2.4). Since by the details between (2.2) and (2.4), $\varphi^*_n(x) = \varphi^\circ_n(\lceil ak_n \rceil / k_n) - \varphi^\circ_n(x)$, $\lceil ak_n \rceil / k_n \leqq x \leqq$ $\leqq \lceil bk_n \rceil / k_n$, and $\varphi^*_n(x) = \varphi^\circ_n(\lceil ak_n \rceil / k_n) - \varphi^\circ_n(\lceil bk_n \rceil / k_n)$, $x \geqq \lceil bk_n \rceil / k_n$, for the functions $\varphi^\circ_n(x) = 2^{-\lceil l_n + \operatorname{Log} x \rceil}$, $x > 0$, in Lemma 2, we have $\varphi^*_n(\lceil bk_n \rceil / k_n) =$ $= \varphi^\circ_n(\lceil ak_n \rceil / k_n) - \varphi^\circ_n(\lceil bk_n \rceil / k_n) = \Delta^\circ_n(a,b)$. Hence we see that $\varphi^*_n \not\equiv 0$ if and only if $\Delta^\circ_n(a,b) > 0$, which happens if and only if $\lceil \operatorname{Log} b \rceil - \lceil \operatorname{Log} a \rceil + \lfloor \gamma_a -$ $- z_n(a) \rfloor - \lfloor \gamma_b - z_n(b) \rfloor \geqq 1$, and so, recalling $\kappa = \kappa_{a,b} = \lceil \operatorname{Log} b \rceil - \lceil \operatorname{Log} a \rceil \in$ $\in \{0,1,2,\dots\}$ from (1.10), noting again that $\kappa = 0$ implies $\gamma_a > \gamma_b$,

$$\varphi^*_n \not\equiv 0 \quad \text{if and only if} \quad \kappa + \lfloor \gamma_a - z_n(a) \rfloor - \lfloor \gamma_b - z_n(b) \rfloor \geqq 1. \tag{2.13}$$

Recall also $\varphi^\circ_l(x) = 2^{-\lceil l + \operatorname{Log} x \rceil}$, $x > 0$, from Lemma 2.

LEMMA 5. *Let $\{k_{n'}\}$ be positive numbers such that* (1.1) *is satisfied along $\{n'\} \subset \mathbf{N}$. Suppose that $l_{n'} \to l$ and $\varphi^*_{n'}(\cdot) \Longrightarrow \varphi^*(\cdot)$ on $\mathbf{R}$, as $n' \to \infty$, for some $l \in [0,1]$ and some function $\varphi^*(\cdot) \not\equiv 0$. Then there exists a point $x_0 \in$ $\in [a,b]$, at which the functions $\varphi^*(\cdot)$ and $\varphi^\circ_l(\cdot)$ jump simultaneously, such that $\gamma_{x_0} = l$ if $l < 1$ and $\gamma_{x_0} = 0$ if $l = 1$.*

PROOF. Under the stated conditions, the inequality in (2.13) holds for all n' large enough. The function φ^* is necessarily non-decreasing on $\mathbf{R}$ and we take it to be left-continuous. Since $\varphi^\circ_{n'}(\lceil ck_{n'}\rceil/k_{n'}) = 2^{-\lceil \operatorname{Log} c\rceil + \lfloor \gamma_c - z_{n'}(c)\rfloor}$, $c = a, b$, an application of Lemma 2 gives

$$\varphi^*(x) = \begin{cases} 0, & \text{if } x \leqq a, \\ 2^{-\lceil \operatorname{Log} a\rceil + m'_a} - 2^{-\lceil l + \operatorname{Log} x\rceil}, & \text{if } a < x \leqq b, \\ 2^{-\lceil \operatorname{Log} a\rceil + m'_a} - 2^{-\lceil \operatorname{Log} b\rceil + m'_b}, & \text{if } b < x, \end{cases} \tag{2.14}$$

where $m'_c := \lim_{n'\to\infty} \lfloor \gamma_c - z_{n'}(c)\rfloor \in \{-2,-1,0\}$ necessarily exists for both $c = a$ and $c = b$. Note that by (2.1), $m_c(n') := \lfloor \gamma_c - z_{n'}(c)\rfloor = \lfloor \gamma_c - l_{n'} - \varepsilon_{n'}(c)\rfloor$, where $\varepsilon_{n'}(c) := \operatorname{Log}(\lceil ck_{n'}\rceil/(ck_{n'})) \geqq 0$ and $\lim_{n'\to\infty} \varepsilon_{n'}(c) = 0$, so that $\lim_{n'\to\infty} m_c(n') = m'_c$ if and only if $m_c(n') = m'_c$ for all n' large enough, $c = a, b$. Furthermore, either $m'_c = \lfloor \gamma_c - l\rfloor$ or $m'_c = \lfloor \gamma_c - l\rfloor - 1$, $c = a, b$, but from (2.13) we obtain

$$\varphi^*(\cdot) \not\equiv 0 \quad \text{if and only if} \quad \kappa + m'_a - m'_b \geqq 1. \tag{2.15}$$

(This also follows directly, using the criterion $\varphi^*(b+) > 0$ in (2.14) itself.) The discontinuity points of φ^* and φ°_l are the same in (a, b). From the proof of Lemma 2, the set of jump-points of φ°_l in $[a, b]$ is nothing but $\{a \leqq x \leqq b : l + \operatorname{Log} x \in \mathbf{Z}\} = \{a \leqq x \leqq b : \gamma_x - l \in \mathbf{Z}\}$.

The statement is clear, with some $x_0 \in (a, b)$, if $\operatorname{Log} b - \operatorname{Log} a > 1$. (Because $b > 2a$, if $l = 0, 1$ or $0 < l < 1$, the respective choices $a < x_0 = 2^j < b$ or $a < x_0 = 2^{j-l} < b$ are always possible for some $j \in \mathbf{Z}$.) Since $\operatorname{Log} b - \operatorname{Log} a = \lceil \operatorname{Log} b\rceil - \gamma_b - \lceil \operatorname{Log} a\rceil + \gamma_a = \kappa + (\gamma_a - \gamma_b)$, this is satisfied whenever $\kappa \geqq 2$, i.e. if any one of conditions $\mathbf{C}_1, \dots, \mathbf{C}_{17}$ is satisfied. (The seemingly contradictory case when $\kappa = 2$ and $(m'_a, m'_b) = (-2, 0)$, which would force $\varphi^*(\cdot) \equiv 0$ by (2.15), cannot happen: If $m'_a = -2$, then necessarily $l = 1$ and $\gamma_a = 0$, which since $\gamma_b - 1 < 0$ implies $m'_b \leqq -1$; the real cases that arise are covered in $\mathbf{C}_6$ and $\mathbf{C}_9$.) It is also satisfied if $\kappa = 1$ and $\gamma_a > \gamma_b$, i.e. if we have condition $\mathbf{C}_{18}$.

Next consider the case $\mathbf{C}_{19} : \kappa = 1,\ \gamma_a = \gamma_b =: \gamma$. Note that necessarily $b = 2a$, and using the elementary inequality that $\lceil 2x\rceil \geqq 2\lceil x\rceil$, $x \in \mathbf{R}$, we have $\lceil bk_{n'}\rceil \geqq 2\lceil ak_{n'}\rceil$, so that $\lceil bk_{n'}\rceil/bk_{n'} \geqq \lceil ak_{n'}\rceil/ak_{n'}$, and hence $\lfloor \gamma_a - z_{n'}(a)\rfloor \geqq \lfloor \gamma_b - z_{n'}(b)\rfloor$, since $\gamma_a = \gamma_b$. So, $\mathbf{C}_{19}$ implies (2.15), thus it is compatible with $\varphi^*(\cdot) \not\equiv 0$. Therefore, in particular, we see by (2.15) that $(m'_a, m'_b) \neq (-2, -1)$ and $(m'_a, m'_b) \neq (-1, 0)$ in this case.

If $l < \gamma$, then $l + \operatorname{Log} a < \gamma + \operatorname{Log} a = \lceil \operatorname{Log} a\rceil = \lceil \operatorname{Log} b\rceil - 1 < \lceil \operatorname{Log} b\rceil - \gamma = \operatorname{Log} b \leqq l + \operatorname{Log} b$, and so $\operatorname{Log} x_0 = \lceil \operatorname{Log} a\rceil - l \in (\operatorname{Log} a, \operatorname{Log} b)$. If $\gamma < l < 1$ or $0 < \gamma < l = 1$, then $l + \operatorname{Log} b > \gamma + \operatorname{Log} b = \lceil \operatorname{Log} b\rceil = \lceil \operatorname{Log} a\rceil + 1 > \lceil \operatorname{Log} a\rceil + l - \gamma = l + \operatorname{Log} a$, so that $\operatorname{Log} x_0 = \lceil \operatorname{Log} b\rceil - l \in (\operatorname{Log} a, \operatorname{Log} b)$. Two subcases of $\mathbf{C}_{19}$ remain, the first is $\mathbf{C}^1_{19} : l = 1, \gamma = 0$, while the second is $\mathbf{C}^2_{19} : l = \gamma$.

Under $\mathbf{C}^1_{19}$ we either have $(m'_a, m'_b) = (-2, -2)$ in (2.14), in which case $x_0 = b$; or $(m'_a, m'_b) = (-1, -1)$, in which case $x_0 = a$; or $(m'_a, m'_b) =$

$= (-1,-2)$, in which case both $x_0 = a$ and $x_0 = b$ work. Under $\mathbf{C}_{19}^2$ we either have $(m'_a, m'_b) = (0,0)$, in which case $x_0 = a$; or $(m'_a, m'_b) = (-1,-1)$, in which case $x_0 = b$; or $(m'_a, m'_b) = (0,-1)$, in which case both $x_0 = a$ and $x_0 = b$ are possible choices.

Some cases will be ruled out from the present considerations since $\varphi^*(\cdot) \equiv 0$. However, listing them here will be useful later on. These cases are

$$\begin{aligned} &\mathbf{C}_{29}^0 : \kappa = 1, \gamma_a < l < \gamma_b, \text{ when necessarily } (m'_a, m'_b) = (-1,0), \\ &\mathbf{C}_{30}^0 : \kappa = 0, \gamma_b < \gamma_a < l, \text{ when necessarily } (m'_a, m'_b) = (-1,-1), \qquad (2.16) \\ &\mathbf{C}_{31}^0 : \kappa = 0, l < \gamma_b < \gamma_a, \text{ when necessarily } (m'_a, m'_b) = (0,0), \end{aligned}$$

so that the inequality in (2.15) fails, for all n' large enough, in all three cases.

Resuming the proof, consider $\mathbf{C}_{20} : \kappa = 1,\ l < \gamma_a < \gamma_b$. Then $l + \operatorname{Log} a < < \gamma_a + \operatorname{Log} a = \lceil \operatorname{Log} a \rceil = \lceil \operatorname{Log} b \rceil - 1 < \lceil \operatorname{Log} b \rceil - (\gamma_b - l) = l + \operatorname{Log} b$, so that $\operatorname{Log} x_0 = \lceil \operatorname{Log} a \rceil - l \in (\operatorname{Log} a, \operatorname{Log} b)$.

Next, for $\mathbf{C}_{21} : \kappa = 1,\ l = \gamma_a < \gamma_b$, we have $m'_b = 0$, and so by (2.15), $m'_a = 0$. Thus, substituting these values into (2.14), we see that $x_0 = a$.

If $\mathbf{C}_{22} : \kappa = 1,\ \gamma_a < \gamma_b = l$, holds, then $m'_a = -1$, which by (2.15) forces $m'_b = -1$, so that from (2.14) we have $x_0 = b$.

Under $\mathbf{C}_{23} : \kappa = 1,\ \gamma_a < \gamma_b < l < 1$, or $0 < \gamma_a < \gamma_b < l = 1$, we have $l + \operatorname{Log} b > \lceil \operatorname{Log} b \rceil = \lceil \operatorname{Log} a \rceil + 1 = \gamma_a + 1 + \operatorname{Log} a > l + \operatorname{Log} a$, so $\operatorname{Log} x_0 = = \lceil \operatorname{Log} b \rceil - l \in (\operatorname{Log} a, \operatorname{Log} b)$.

For $\mathbf{C}_{24} : \kappa = 1,\ 0 = \gamma_a < \gamma_b < l = 1$, we must have $m'_b = -1$, and thus by (2.15) also $m'_a = -1$, so from (2.14), $x_0 = a$ again.

Finally, the cases with $\kappa = 0$ follow, two of which, with $\gamma_b < \gamma_a$, have already been excluded in (2.16). When $\mathbf{C}_{25} : \kappa = 0,\ l = \gamma_b < \gamma_a$, holds, then $(m'_a, m'_b) = (0,-1)$, using also (2.15) to force the second coordinate, so that from (2.14), $x_0 = b$.

If $\mathbf{C}_{26} : \kappa = 0,\ \gamma_b < l < \gamma_a$, obtains, then $l + \operatorname{Log} b > \gamma_b + \operatorname{Log} b = \lceil \operatorname{Log} b \rceil = = \lceil \operatorname{Log} a \rceil = \gamma_a + \operatorname{Log} a > l + \operatorname{Log} a$, so $\operatorname{Log} x_0 = \lceil \operatorname{Log} a \rceil - l \in (\operatorname{Log} a, \operatorname{Log} b)$.

Next, for $\mathbf{C}_{27} : \kappa = 0,\ \gamma_b < \gamma_a = l$, necessarily $m'_b = -1$, which by (2.15) forces $m'_a = 0$, so that from (2.14), $x_0 = a$.

Lastly, in the case of $\mathbf{C}_{28} : \kappa = 0,\ 0 = \gamma_b < \gamma_a,\ l = 1$, we must have $m'_a = = -1$, and hence by (2.15) also $m'_b = -2$, which by (2.14) yields the choice $x_0 = b$. □

Lemma 6. *Under the conditions of Lemma 5, if $\varphi^*(\cdot)$ jumps at $x \in [a,b]$, then either $\gamma_x = l$, in which case $\varphi^*(x+) - \varphi^*(x) = 2^{-\lceil \operatorname{Log} x \rceil - 1}$, or $\gamma_x = 0$ and $l = 1$, in which case $\varphi^*(x+) - \varphi^*(x) = 2^{-\lceil \operatorname{Log} x \rceil - 2}$.*

Proof. If the jump-point $x \in (a,b)$, then $\varphi^\circ(\cdot)$ jumps at x, and the stated alternative is trivial. If, on the other hand, the jump-point is either $x = a$ or $x = b$ and the stated alternative fails to hold, then for $c = a, b$ we have $-\lceil l + \operatorname{Log} x \rceil \to -\lceil l + \operatorname{Log} c \rceil = -\lceil \operatorname{Log} c \rceil + \lfloor \gamma_c - l \rfloor$ as $x \to c$, and it is clear from (2.14) that $m'_c = \lfloor \gamma_c - l \rfloor - 1$. But since $\gamma_c - z_{n'} \to \gamma_c - l$ as $n' \to \infty$, and $\gamma_c - l \neq 0$, $\gamma_c - l \neq -1$, we also have $m'_c = \lfloor \gamma_c - l \rfloor$, $c = a, b$. Thus, in the endpoints, the desired alternative follows by contradiction. The statements concerning the saltus of φ^* at an inner point $x \in (a, b)$ are trivial by (2.14).

Let $x = a$. If $\gamma_a = l$, then $\lceil l + \operatorname{Log} x \rceil \to \lceil \operatorname{Log} a \rceil + 1$ as $x \downarrow a$, and m'_a in (2.14) is either 0 or -1. But since φ^* jumps at a, we must have $m'_a = 0$, so, the saltus is $2^{-\lceil \operatorname{Log} a \rceil} - 2^{-\lceil \operatorname{Log} a \rceil - 1} = 2^{-\lceil \operatorname{Log} a \rceil - 1}$. If $\gamma_a = 0$ and $l = 1$, then $\lceil l + \operatorname{Log} x \rceil \to \lceil \operatorname{Log} a \rceil + 2$ as $x \downarrow a$, and the possible values of m'_a are -1 or -2. But since φ^* jumps at a, we have $m'_a = -1$, so, the saltus is $2^{-\lceil \operatorname{Log} a \rceil - 2}$. Let $x = b$. If $\gamma_b = l$, then $\lceil l + \operatorname{Log} x \rceil \to \lceil \operatorname{Log} b \rceil$ as $x \uparrow b$, and m'_b is either 0 or -1, so we must have $m_b = -1$; thus the saltus is $2^{-\lceil \operatorname{Log} b \rceil} - 2^{-\lceil \operatorname{Log} b \rceil - 1} = 2^{-\lceil \operatorname{Log} b \rceil - 1}$. If $\gamma_b = 0$ and $l = 1$, then $\lceil l + \operatorname{Log} x \rceil \to \lceil \operatorname{Log} b \rceil + 1$ as $x \uparrow b$, and m'_b is either -1 or -2, so we have $m'_b = -2$; thus the saltus is $2^{-\lceil \operatorname{Log} b \rceil - 1} - 2^{-\lceil \operatorname{Log} b \rceil - 2} = 2^{-\lceil \operatorname{Log} b \rceil - 2}$. □

To sort out which cases yield $\varphi^*(\cdot) \not\equiv 0$ and $\varphi^*(\cdot) \equiv 0$ if $\varphi^*_{n'}(\cdot) \Longrightarrow \varphi^*(\cdot)$ on $\mathbf{R}$ along some $\{n'\} \subset \mathbf{N}$, we need the following additional shorthand:

$$\begin{aligned}
&\mathbf{C}^+_{21} : \mathbf{C}_{21} \text{ holds and } (m'_a, m'_b) = (0, 0),\\
&\mathbf{C}^+_{22} : \mathbf{C}_{22} \text{ holds and } (m'_a, m'_b) = (-1, -1)\\
&\mathbf{C}^+_{24} : \mathbf{C}_{24} \text{ holds and } (m'_a, m'_b) = (-1, -1),\\
&\mathbf{C}^+_{25} : \mathbf{C}_{25} \text{ holds and } (m'_a, m'_b) = (0, -1),\\
&\mathbf{C}^+_{27} : \mathbf{C}_{27} \text{ holds and } (m'_a, m'_b) = (0, -1),\\
&\mathbf{C}^+_{28} : \mathbf{C}_{28} \text{ holds and } (m'_a, m'_b) = (-1, -2);
\end{aligned}$$

$$\begin{aligned}
&\mathbf{C}^0_{21} : \mathbf{C}_{21} \text{ holds and } (m'_a, m'_b) = (-1, 0),\\
&\mathbf{C}^0_{22} : \mathbf{C}_{22} \text{ holds and } (m'_a, m'_b) = (-1, 0),\\
&\mathbf{C}^0_{24} : \mathbf{C}_{24} \text{ holds and } (m'_a, m'_b) = (-2, -1),\\
&\mathbf{C}^0_{25} : \mathbf{C}_{25} \text{ holds and } (m'_a, m'_b) = (0, 0),\\
&\mathbf{C}^0_{27} : \mathbf{C}_{27} \text{ holds and } (m'_a, m'_b) = (-1, -1),\\
&\mathbf{C}^0_{28} : \mathbf{C}_{28} \text{ holds and } (m'_a, m'_b) = (-1, -1).
\end{aligned}$$

Recall also the notation $\mathbf{C}^0_{29}, \mathbf{C}^0_{30}, \mathbf{C}^0_{31}$ in (2.16).

LEMMA 7. *Let $\{k_{n'}\}$ be positive numbers such that* (1.1) *is satisfied along $\{n'\} \subset \mathbf{N}$.*

(1) *If $\varphi^*_{n'}(\cdot) \Longrightarrow \varphi^*(\cdot) \not\equiv 0$ on $\mathbf{R}$ as $n' \to \infty$, then $\lim_{n' \to \infty} l_{n'} = l$ for some $l \in [0, 1]$.*

(2) *If $\varphi^*_{n'}(\cdot) \Longrightarrow \varphi^*(\cdot)$ on $\mathbf{R}$ and $l_{n'} \to l$ as $n' \to \infty$, then $\varphi^*(\cdot) \not\equiv 0$ if and only if the quintuplet $(\kappa_{a,b}, \gamma_a, \gamma_b, l, \{n'\})$ satisfies one of the twenty-eight conditions $\mathbf{C}_1, \ldots, \mathbf{C}_{20}$, $\mathbf{C}^+_{21}$, $\mathbf{C}^+_{22}$, $\mathbf{C}_{23}$, $\mathbf{C}^+_{24}$, $\mathbf{C}^+_{25}$, $\mathbf{C}_{26}$, $\mathbf{C}^+_{27}$, $\mathbf{C}^+_{28}$.*

(3) *If* $\varphi^*_{n'}(\cdot) \Longrightarrow \varphi^*(\cdot)$ *on* $\mathbf{R}$ *and* $l_{n'} \to l$ *as* $n' \to \infty$, *then* $\varphi^*(\cdot) \equiv 0$ *if and only if the quintuplet* $(\kappa_{a,b}, \gamma_a, \gamma_b, l, \{n'\})$ *satisfies one of the nine conditions* $\mathbf{C}^0_{21}$, $\mathbf{C}^0_{22}$, $\mathbf{C}^0_{24}$, $\mathbf{C}^0_{25}$, $\mathbf{C}^0_{27}$, $\mathbf{C}^0_{28}$, $\mathbf{C}^0_{29}$, $\mathbf{C}^0_{30}$, $\mathbf{C}^0_{31}$.

(4) *Suppose that* $\varphi^*_{n'}(\cdot) \Longrightarrow \varphi^*(\cdot)$ *on* $\mathbf{R}$ *and* $l_{n'} \to l$ *as* $n' \to \infty$. *If* $\mathbf{C}^0_{29}$ *holds, then necessarily condition* $a_1 \sqcap b_2$ *is satisfied along* $\{n'\} \subset \mathbf{N}$; *if* $\mathbf{C}^0_{30}$ *holds, then condition* $a_1 \sqcap b_1$ *is satisfied; while if* $\mathbf{C}^0_{31}$ *holds, then condition* $a_2 \sqcap b_2$ *is satisfied. In all three cases,* $(\psi^*_{n'}(a;\cdot), \varphi^*_{n'}(\cdot), \psi^*_{n'}(b;\cdot)) \Longrightarrow (\psi^*_0(\cdot), \varphi^*(\cdot), \psi^*_0(\cdot)) \equiv (0,0,0)$ *as* $n' \to \infty$.

PROOF. (1) Choose two subsequences $\{n''\} \subset \{n'\}$ and $\{n'''\} \subset \{n'\}$ such that $\lim_{n''\to\infty} l_{n''} = l''$ and $\lim_{n'''\to\infty} l_{n'''} = l'''$. Then by Lemma 2, $\varphi^\circ_{n''}(\cdot) \Longrightarrow \varphi^\circ_{l''}(\cdot)$ as $n'' \to \infty$ and $\varphi^\circ_{n'''}(\cdot) \Longrightarrow \varphi^\circ_{l'''}(\cdot)$ as $n''' \to \infty$, on $(0,\infty)$. If at least one of $\varphi^\circ_{l''}$ and $\varphi^\circ_{l'''}$ jumps in (a,b), so does φ^* by (2.14). By Lemma 5, either $l'' = l'''$, or $\gamma_x = 0 = l''$ and $l''' = 1$, or $\gamma_x = 0 = l'''$ and $l'' = 1$. But Lemma 6 rules out the last two possibilities, and so $l'' = l'''$. If neither $\varphi^\circ_{l''}$ nor $\varphi^\circ_{l'''}$ jumps in (a,b), then φ^* cannot jump in (a,b) either, but then φ^* and $\varphi^\circ_{l''}$ jump simultaneously at x'' and φ^* and $\varphi^\circ_{l'''}$ jump simultaneously at x''', where both x'' and x''' are one of the endpoints of $[a,b]$, or both. If $x'' = c = x'''$, $c = a$ or $c = b$, then, by Lemma 6, either $l'' = \gamma_c = l'''$, or $\gamma_c = 0$ and $l'' = 1 = l'''$. If $x'' \neq x'''$, for definiteness $x'' = a$ and $x''' = b$, say, then, by the proof of Lemma 5, either $(m''_a, m''_b) = (-1,-2)$ or $(m'''_a, m'''_b) = (0,-1)$ in (2.14), where the superscripts $''$ or $'''$ refer to the underlying subsequence. In the first case, $l'' = 1 = l'''$, while in the second, $l'' = \gamma_a = \gamma_b = l'''$. Thus, $\lim_{n'\to\infty} l_{n'}$ exists.

(2&3) Both statements here follow from a detailed review of the proof of Lemma 5, in combination with (2.15).

(4) The statements about the $a_i \sqcap b_j$ conditions follow by obvious direct inspection, while the degeneracy statement for the limiting triplet comes from Lemma 4(1&2) and part (3) here. □

LEMMA 8. *Let* $\{k_{n'}\}$ *be positive numbers such that* (1.1) *is satisfied along* $\{n'\} \subset \mathbf{N}$. *Suppose that* $\lim_{n'\to\infty} l_{n'} = l$ *and* $(\kappa_{a,b}, \gamma_a, \gamma_b, l, \{n'\})$ *satisfies one of* $\mathbf{C}_1, \ldots, \mathbf{C}_{20}$, $\mathbf{C}^+_{21}$, $\mathbf{C}^+_{22}$, $\mathbf{C}_{23}$, $\mathbf{C}^+_{24}$, $\mathbf{C}^+_{25}$, $\mathbf{C}_{26}$, $\mathbf{C}^+_{27}$, $\mathbf{C}^+_{28}$, *and, in addition, that one of the conditions* $a_i \sqcap b_j$, $1 \leqq i, j \leqq 4$, *is also satisfied. Then every subsequence* $\{n''\} \subset \{n'\}$ *contains a further subsequence* $\{n'''\} \subset \{n''\}$ *such that* $\varphi^*_{n'''}(\cdot) \Longrightarrow \varphi^*_l(\cdot)$ *on* $\mathbf{R}$ *as* $n''' \to \infty$, *where* $\varphi^*_l(\cdot) \not\equiv 0$ *and has the form*

$$\varphi^*_l(x) = \begin{cases} 0, & \text{if } x \leqq a, \\ 2^{-\lceil \operatorname{Log} a \rceil + m'''_a} - 2^{-\lceil l + \operatorname{Log} x \rceil}, & \text{if } a < x \leqq b, \\ 2^{-\lceil \operatorname{Log} a \rceil + m'''_a} - 2^{-\lceil \operatorname{Log} b \rceil + m'''_b}, & \text{if } b < x, \end{cases} \tag{2.17}$$

where, for $c=a,b$, the quantities m_c''' are given by

$$m_c'''=\begin{cases}\lambda(c_i), & \textit{if } c_i\notin\{c_3^0,c_4^0\},\\ \lambda(c_3)=0, & \textit{if } c_3^0 \textit{ holds and } c_3^{0+} \textit{ obtains along } \{n'''\},\\ \lambda^*(c_3^{0-})=-1, & \textit{if } c_3^0 \textit{ holds and } c_3^{0-} \textit{ obtains along } \{n'''\},\\ \lambda(c_4^0)=-1, & \textit{if } c_4^0 \textit{ holds and } c_4^{0+} \textit{ obtains along } \{n'''\},\\ \lambda^*(c_4^{0-})=-2, & \textit{if } c_4^0 \textit{ holds and } c_4^{0-} \textit{ obtains along } \{n'''\}.\end{cases}$$

PROOF. By Lemma 2, $\varphi_{n'}^\circ(\cdot)\Longrightarrow\varphi_l^\circ(\cdot)$ on $(0,\infty)$, and the existence of $\{n'''\}$ with all the stated properties follows by combining Lemma 3 and Lemma 7(2) with the beginning of the proof of Lemma 5. □

Now let $\{W(t):t\geqq 0\}$ be a standard Wiener process, and for $i,j=1,\dots,4$, define

$$J_{i,j}:=J_{i,j}(a,b,l):=\frac{1}{2}\sum_{k=\lceil\operatorname{Log}a\rceil-\lambda(a_i)}^{\lceil\operatorname{Log}b\rceil-\lambda(b_j)-1}\frac{W(2^{k-l})}{2^k},$$

half of the middle sum in (1.12).

LEMMA 9. *If the conditions of Lemma 8 hold and $\varphi_l^*(\cdot)$ is the function in (2.17), then*

$$\int_{[a,b]}W(x)d\varphi_l^*(x)=$$

$$=\begin{cases}J_{i,j}, & \textit{if } (m_a''',m_b''')=(\lambda(a_i),\lambda(b_j)),\\ \frac{-W(a)}{2^{\lceil\operatorname{Log}a\rceil+1}}+J_{3,j}, & \textit{if } (m_a''',m_b''')=(\lambda^*(a_3^{0-}),\lambda(b_j))=(-1,\lambda(b_j)),\\ J_{i,3}+\frac{W(b)}{2^{\lceil\operatorname{Log}b\rceil+1}}, & \textit{if } (m_a''',m_b''')=(\lambda(a_i),\lambda^*(b_3^{0-}))=(\lambda(a_i),-1),\\ \frac{-W(a)}{2^{\lceil\operatorname{Log}a\rceil+2}}+J_{4,j}, & \textit{if } (m_a''',m_b''')=(\lambda^*(a_4^{0-}),\lambda(b_j))=(-2,\lambda(b_j)),\\ J_{i,4}+\frac{W(b)}{2^{\lceil\operatorname{Log}b\rceil+2}}, & \textit{if } (m_a''',m_b''')=(\lambda(a_i),\lambda^*(b_4^{0-}))=(\lambda(a_i),-2),\end{cases}$$

for all $i,j=1,\dots,4$, where the second, …, fifth branches are meant as $a_3^0\sqcap b_j$ holds and a_3^{0-} obtains, $a_i\sqcap b_3^0$ holds and b_3^{0-} obtains, $a_4^0\sqcap b_j$ holds and a_4^{0-} obtains, and $a_i\sqcap b_4^0$ holds and b_4^{0-} obtains along $\{n'''\}$.

PROOF. Throughout, the integral $\int_{[a,b]}W(x)d\varphi_l^*(x)$ in question will be denoted by $I_{i,j}$, suppressing the dependence on other parameters in the notation. Setting $\delta_l(c) := \varphi_l^*(c+) - \varphi_l^*(c)$ for the saltus of $\varphi_l^*(\cdot)$ at $c = a, b$,

$$
\begin{aligned}
I_{i,j} &= \delta_l(a)W(a) + \int_{(a,b)} W(x)\,d\varphi_l^*(x) + \delta_l(b)W(b) = \\
&= \delta_l(a)W(a) + \sum_{k\in(l+\operatorname{Log} a,\, l+\operatorname{Log} b)\cap\mathbf{Z}} \frac{W(2^{k-l})}{2^{k+1}} + \delta_l(b)W(b),
\end{aligned}
\tag{2.18}
$$

where, if $\delta_l(c) > 0$, then by Lemma 6, either $\gamma_c = l$, or $\gamma_c = 0$ and $l = 1$, that is, $l + \operatorname{Log} c \in \mathbf{Z}$ for $c = a, b$. Assuming that $(m_a''', m_b''') = (\lambda(a_i), \lambda(b_j))$ for some $i, j = 1, \ldots, 4$ in (2.17), consider first the first branch of the statement.

If $\gamma_a = l$, then the smallest k in the sum of (2.18) is $\underline{k} = l + 1 + \operatorname{Log} a = \lceil \operatorname{Log} a \rceil + 1$. If $\lambda(a_i) = -1$, then, by (2.17), $\delta_l(a) = 0$, so that, using the convention that $k \in \mathbf{Z}$, (2.18) becomes $I_{i,j} = \sum_{\underline{k} \leqq k < l+\operatorname{Log} b} w_k + \delta_l(b)W(b)$, where $w_k := w_k(l) := W(2^{k-l})/2^{k+1}$ throughout and where $\underline{k} = \lceil \operatorname{Log} a \rceil + 1 = \lceil \operatorname{Log} a \rceil - \lambda(a_i)$. If $\lambda(a_i) = 0$, then, again by (2.17), $\delta_l(a)W(a) = W(2^{\lceil \operatorname{Log} a \rceil + l})/2^{\lceil \operatorname{Log} a \rceil + 1}$, so $I_{i,j} = \sum_{\underline{k}-1 \leqq k < l+\operatorname{Log} b} w_k + \delta_l(b)W(b)$ for all $j = 1, \ldots, 4$, where $\underline{k} - 1 = l + \operatorname{Log} a = \lceil \operatorname{Log} a \rceil = \lceil \operatorname{Log} a \rceil - \lambda(a_i)$.

Now let $\gamma_a = 0$ and $l = 1$. Then the summation for k in (2.18) begins on $\underline{k} = l + 1 + \operatorname{Log} a = \lceil \operatorname{Log} a \rceil + 2$. In this case, $\lambda(a_i) = -1$ necessarily, and, using (2.17), $W(a)\delta_l(a) = W(2^{\lceil \operatorname{Log} a \rceil + 1 - 1})/2^{\lceil \operatorname{Log} a \rceil + 2}$, so $I_{i,j} = \sum_{\underline{k}-1 \leqq k < l+\operatorname{Log} b} w_k + \delta_l(b)W(b)$, where $\underline{k} - 1 = \lceil \operatorname{Log} a \rceil + 1 = \lceil \operatorname{Log} a \rceil - \lambda(a_i)$, again, for all $j = 1, \ldots, 4$. Thus we know that the lower bound of k in $J_{i,j}$ is as stated for the first branch whenever $l + \operatorname{Log} a \in \mathbf{Z}$.

Suppose next that $l + \operatorname{Log} a \notin \mathbf{Z}$. This means that either $l \neq \gamma_a$, or $\gamma_a \neq 0$ when $l = 1$. Then either $0 < l - \gamma_a < 1$, when a_1 holds; or $l < \gamma_a$, when a_2 takes place. The sum in (2.18) begins on $\lceil + \operatorname{Log} a \rceil = \lceil \operatorname{Log} a \rceil + \lceil -\gamma_a + l \rceil = \lceil \operatorname{Log} a \rceil - \lfloor \gamma_a - l \rfloor = \lceil \operatorname{Log} a \rceil - \lambda(a_i)$, and since $\delta_l(a) = 0$, we have $I_{i,j} = \sum_{\lceil \operatorname{Log} a \rceil - \lambda(a_i) \leqq k < l+\operatorname{Log} b} w_k + \delta_l(b)W(b)$ for all $j = 1, \ldots, 4$ and both possible values $i = 1, 2$. Therefore, the lower bound of k in $J_{i,j}$ is always as desired for the first branch.

Now we turn to the upper bound, again starting with integer values of $l + \operatorname{Log} b$. First let $l = \gamma_b$. Then the summation in (2.18) ends on $\overline{k} = l - 1 + \operatorname{Log} b = \lceil \operatorname{Log} b \rceil - 1$. If $\lambda(b_j) = 0$, then, as (2.17) shows, $\delta_l(b) = 0$, and so $I_{i,j} = \sum_{\lceil \operatorname{Log} a \rceil - \lambda(a_i) \leqq k \leqq \overline{k}} w_k$, where $\overline{k} = \lceil \operatorname{Log} b \rceil - 1 = \lceil \operatorname{Log} b \rceil - \lambda(b_j) - 1$. Likewise, if $\lambda(b_j) = -1$, then $\delta_l(b)W(b) = W(2^{\lceil \operatorname{Log} b \rceil})/2^{\lceil \operatorname{Log} b \rceil + 1}$ by (2.17), and so $I_{i,j} = \sum_{\lceil \operatorname{Log} a \rceil - \lambda(a_i) \leqq k \leqq \overline{k}+1} w_k$ for all $i = 1, \ldots, 4$, where $\overline{k} + 1 = \lceil \operatorname{Log} b \rceil = \lceil \operatorname{Log} b \rceil - \lambda(b_j) - 1$.

Next, let $\gamma_b = 0$ and $l = 1$. In this case, $\lambda(b_j) = -1$ necessarily, and $\delta_l(b) = 0$ by (2.17). The sum in (2.18) ends on $\overline{k} = l - 1 + \operatorname{Log} b = \operatorname{Log} b = \lceil \operatorname{Log} b \rceil - \lambda(b_j) - 1$.

Finally, let $l + \operatorname{Log} b \notin \mathbf{Z}$. Then $\delta_l(b) = 0$, and either $l \neq \gamma_b$, or $\gamma_b \neq 0$ when $l = 1$. This means that either $0 < l - \gamma_b < 1$, and then b_1 holds; or $l < \gamma_b$, when b_2 takes place. In both cases the summation in (2.18) ends on $\lfloor l + \operatorname{Log} b \rfloor = \lceil \operatorname{Log} b \rceil + \lfloor -\gamma_b + l \rfloor = \lceil \operatorname{Log} b \rceil - 1 + \lceil -\gamma_b + l \rceil = \lceil \operatorname{Log} b \rceil - 1 - \lfloor \gamma_b - l \rfloor = \lceil \operatorname{Log} b \rceil - \lambda(b_j) - 1$. Thus the first branch of the lemma is fully established.

Assuming $a_3^0 \sqcap b_j$ and a_3^{0-} along $\{n'''\}$, so that $(m_a''', m_b''') = (\lambda^*(a_3^{0-}), \lambda(b_j)) = (-1, \lambda(b_j))$ for some $j = 1, \dots, 4$ in (2.17), consider the second branch. By Lemma 3(3), $l = \gamma_a$ and $\delta_l(a) = 0$ as (2.17) shows. In this case, the sum in (2.18) commences on $\underline{k} = l + 1 + \operatorname{Log} a = \lceil \operatorname{Log} a \rceil + 1$ and, since $\underline{k} - 1 = \lceil \operatorname{Log} a \rceil = \lceil \operatorname{Log} a \rceil - \lambda(a_3^0)$, we have $I_{3,j} = \sum_{\underline{k}-1 \leqq k \leqq \lceil \operatorname{Log} b \rceil - \lambda(b_j) - 1} w_k - W(2^{\operatorname{Log} a}) 2^{-\lceil \operatorname{Log} a \rceil - 1} = J_{3,j} - W(a) 2^{-\lceil \operatorname{Log} a \rceil - 1}$.

Now suppose $a_i \sqcap b_3^0$, b_3^{0-} along $\{n'''\}$, and so $(m_a''', m_b''') = (\lambda(a_i), \lambda^*(b_3^{0-})) = (\lambda(a_i), -1)$ for some $i = 1, \dots, 4$ in (2.17). Then $l = \gamma_b$ by Lemma 3(3), $\delta_l(b) = 2^{-\lceil \operatorname{Log} b \rceil - 1}$, the greatest value of k in the sum in (2.18) is $\overline{k} = l - 1 + \operatorname{Log} b = \lceil \operatorname{Log} b \rceil - 1 = \lceil \operatorname{Log} b \rceil - \lambda(b_3) - 1$, and thus $I_{i,3} = J_{i,3} + W(b) 2^{-\lceil \operatorname{Log} b \rceil - 1}$ for the third branch.

Next, having $a_4^0 \sqcap b_j$, a_4^{0-}, $(m_a''', m_b''') = (\lambda^*(a_4^{0-}), \lambda(b_j)) = (-2, \lambda(b_j))$ for some $j = 1, \dots, 4$ in (2.17), $\gamma_a = 0$, $l = 1$ from Lemma 3(4), and so $\delta_l(a) = 0$, $\underline{k} = l + 1 + \operatorname{Log} a = \lceil \operatorname{Log} a \rceil + 2$. Since $\underline{k} - 1 = \lceil \operatorname{Log} a \rceil - \lambda(a_4)$, for the fourth branch we obtain $I_{4,j} = \sum_{\underline{k}-1 \leqq k \leqq \lceil \operatorname{Log} b \rceil - \lambda(b_j) - 1} w_k - W(2^{\underline{k}-1-l}) 2^{-\lceil \operatorname{Log} a \rceil - 2} = J_{4,j} - W(a) 2^{-\lceil \operatorname{Log} a \rceil - 2}$.

Finally, when $a_i \sqcap b_4^0$, b_4^{0-} and $(m_a''', m_b''') = (\lambda(a_i), \lambda^*(b_4^{0-})) = (\lambda(a_i), -2)$ obtain, $\gamma_b = 0$, $l = 1$ by Lemma 3(4); thus $\delta_l(b) = 2^{-\lceil \operatorname{Log} b \rceil - 2}$, $\overline{k} = l - 1 + \operatorname{Log} b = \lceil \operatorname{Log} b \rceil - \lambda(b_4) - 1$ and $I_{i,4} = J_{i,4} + W(b) 2^{-\lceil \operatorname{Log} b \rceil - 2}$ for each $i = 1, \dots, 4$. □

For three functions $\psi_a(\cdot), \varphi(\cdot), \psi_b(\cdot)$ on $\mathbf{R}$, introduce

$$V(\psi_a, \varphi, \psi_b) :=$$

$$:= \frac{2}{2^{\lceil \operatorname{Log} a \rceil}} \int_0^{-W(a)} \psi_a(x)\,dx + 2 \int_{[a,b]} W(x)\,d\varphi(x) - \frac{2}{2^{\lceil \operatorname{Log} b \rceil}} \int_0^{-W(b)} \psi_b(x)\,dx.$$

The last lemma fits the behavior of the three basic sequences of functions together.

LEMMA 10. *Let $\{k_{n'}\}$ be positive numbers with (1.1) satisfied along $\{n'\} \subset \subset \mathbf{N}$.*

(1) *Suppose that $\psi^*_{n'}(c;\cdot) \Longrightarrow \psi^*_c(\cdot)$, $c = a, b$, and $\varphi^*_{n'}(\cdot) \Longrightarrow \varphi^*(\cdot)$ as $n' \to \infty$ for some limiting functions such that $(\psi^*_a(\cdot), \varphi^*(\cdot), \psi^*_b(\cdot)) \not\equiv (0,0,0)$. Then $l = \lim_{n'\to\infty} l_{n'}$ exists and for some $m=1,\ldots,28$ the quadruplet $(\kappa, \gamma_a, \gamma_b, l)$ satisfies $\mathbf{C}_m$ and one of the conditions $a_i \sqcap b_j$ in the set C_m holds along $\{n'\}$. In this case,*

$$V(\psi^*_a, \varphi^*, \psi^*_b) = V(l, a_i, b_j) \tag{2.19}$$

for the random variable in (1.12).

(2) *If $\lim_{n'\to\infty} l_{n'} = l$ and for some $m = 1,\ldots,28$ the quadruplet $(\kappa, \gamma_a, \gamma_b, l)$ satisfies $\mathbf{C}_m$ and one of the conditions $a_i \sqcap b_j$ in the set C_m holds along $\{n'\}$, then every subsequence $\{n''\} \subset \{n'\}$ contains a further subsequence $\{n'''\} \subset \subset \{n''\}$ such that $\psi^*_{n'''}(c;\cdot) \Longrightarrow \psi^*_c(\cdot)$, $c = a, b$, and $\varphi^*_{n'''}(\cdot) \Longrightarrow \varphi^*(\cdot)$ as $n''' \to \to \infty$ for some limiting functions such that $(\psi^*_a(\cdot), \varphi^*(\cdot), \psi^*_b(\cdot)) \not\equiv (0,0,0)$, and then* (2.19) *holds again.*

PROOF. (1) If $\varphi^* \not\equiv 0$, then $l = \lim_{n'\to\infty} l_{n'}$ exists by Lemma 7(1). If $\varphi^* \equiv 0$, then $\psi^*_c \not\equiv 0$ for at least one value of $c = a, b$, in which case either c_3 or c_4 holds by Lemma 4(II), and both of these imply that $l = \lim_{n'\to\infty} l_{n'}$ exists.

Next we show that one of the couples $a_i \sqcap b_j$, $i, j = 1, \ldots, 4$ must take place along $\{n'\}$. This directly follows by Lemma 4(II) again, with $i, j = 3, 4$, if none of ψ^*_a and ψ^*_b is identically zero.

If $\psi^*_a \equiv 0$ and $\psi^*_b \not\equiv 0$, then either b_3 or b_4 holds along $\{n'\}$, but Lemma 4(II) allows subsequences $\{n''\}, \{n'''\} \subset \{n'\}$ along which a_1 and a_2 hold, respectively. Suppose this is so. Then $\gamma_a = l < 1$ by Lemma 3, which rules b_4 out, thus b_3 must take place, which implies $l = \gamma_b$. Hence both a_2 and b_3 hold along $\{n'''\}$ with $\gamma_a = \gamma_b$, which is impossible. Thus either a_1 or a_2 must hold along the original $\{n'\}$. The reversed case when $\psi^*_a \not\equiv 0$ and $\psi^*_b \equiv 0$ is completely analogous: we must have $a_i \sqcap b_j$ along $\{n'\}$ for $i = 3$ or 4 and $j = 1$ or 2.

Finally, suppose that both $\psi^*_a \equiv 0$ and $\psi^*_b \equiv 0$, when $\varphi^* \not\equiv 0$. By Lemma 4(II), every subsequence of $\{n'\}$ has a further subsequence along which one of $a_1 \sqcap b_1$, $a_1 \sqcap b_2$, $a_2 \sqcap b_1$, $a_2 \sqcap b_2$ is satisfied. If two of these could be obtained along different subsequences, then, using the first statement in Lemma 7(2) and Lemma 3(1&2), it would follow through a detailed inspection of (2.14) that the limiting functions were different along these two subsequences. But they both must agree with φ^*, a contradiction. So, in summary, $l = \lim_{n'\to\infty} l_{n'}$ exists for some $l \in [0,1]$ and one of $a_i \sqcap b_j$ is satisfied along $\{n'\}$.

Using the definitions in (1.9) and (1.10), it is a simple matter to see that the possible relationships that the quadruplet $(\kappa, \gamma_a, \gamma_b, l)$ may satisfy are exhausted by the conditions listed in $\mathbf{C}_1, \ldots, \mathbf{C}_{28}$ and in $\mathbf{C}^0_{29}, \mathbf{C}^0_{30}, \mathbf{C}^0_{31}$, the latter three given in (2.16). But those in $\mathbf{C}_{29}^0, \mathbf{C}_{30}^0, \mathbf{C}_{31}^0$ are ruled out by Lemma 7(4), since they yield a zero function triplet, and so we are left with one of $\mathbf{C}_1, \dots, \mathbf{C}_{28}$. Given any one of these, $\mathbf{C}_m$, it is then straightforward to see that the pair of conditions $a_i \sqcap b_j$ that must hold along $\{n'\}$ is compatible with $\mathbf{C}_m$ if and only if $a_i \sqcap b_j \in \mathcal{C}_m$, $m = 1, \dots, 28$. (Here we also use the fact, noted above (1.10), that the pairs $a_2 \sqcap b_4$, $a_3 \sqcap b_4$, $a_4 \sqcap b_2$ and $a_4 \sqcap b_3$ are impossible from the very beginning.)

To show (2.19), note that if for $c = a$ or $c = b$, or both, $\psi_c^* \not\equiv 0$, then by Lemma 4(II) one of the seven conditions $c_3^{u-}, c_3^{u+}, c_3^{0-}, c_3^{0+}, c_4^{v+}, c_4^{0-}$ and c_4^{0+} listed before Lemma 3 are satisfied along $\{n'\}$, and, by Lemma 3, for all n' large enough we must have $m_c' = \lfloor \gamma_c - z_{n'}(c) \rfloor = \lambda(c_i)$ for the five cases when c_i is one of $c_3^{u-}, c_3^{u+}, c_3^{0+}, c_4^{v+}, c_4^{0+}$ and $m_c' = \lfloor \gamma_c - z_{n'}(c) \rfloor = \lambda^*(c_3^{0-}) = -1$ for c_3^{0-}, while $m_c' = \lfloor \gamma_c - z_{n'}(c) \rfloor = \lambda^*(c_4^{0+}) = -2$ for c_4^{0+}. Putting these facts together with the definition of $V(\psi_a^*, \varphi^*, \psi_b^*)$ and the integral statements in Lemma 4(II) and Lemma 9, the desired equation in (2.19) follows for all $a_i \sqcap b_j \in \mathcal{C}_m$, under $\mathbf{C}_m$, $m = 1, \dots, 28$, for which either $\lfloor \gamma_a - z_{n'}(a) \rfloor = \lambda(a_i)$ for all n' large enough, or $\lfloor \gamma_b - z_{n'}(b) \rfloor = \lambda(b_j)$ for all n' large enough, or both, $i, j = 1, \dots, 4$. This fact leaves only two last cases to be checked: one is $a_3^{0-} \sqcap b_3^{0-}$ in $\mathcal{C}_5, \mathcal{C}_{10}$ and $\mathcal{C}_{19}$, when necessarily $\gamma_a = l = \gamma_b$, with a special consideration for $\mathcal{C}_{19}$, while the other is $a_4^{0-} \sqcap b_4^{0-}$ in $\mathcal{C}_6$ and $\mathcal{C}_{19}$.

In the first of these, $m_c' = \lfloor \gamma_c - z_{n'}(c) \rfloor = -1$ for all large n', $c = a, b$, and $\varphi^*(\cdot) \equiv \varphi_l^*(\cdot)$ by Lemma 8, so that the $\{n'\}$ version of (2.17) implies $\delta_l(a) = 0$ and $\delta_l(b) = 2^{-\lceil \operatorname{Log} b \rceil - 1}$ for the jumps of $\varphi_l^*(\cdot)$, as used in the proof of Lemma 9. Hence, since for $x \in (a, b)$ the function $\varphi_l^*(\cdot)$ has a saltus $2^{-\lceil \operatorname{Log} x \rceil - 1}$ only if $\gamma_x = \gamma_a = \gamma_b = l$ as seen in Lemma 6, we have

$$2 \int_{[a,b]} W(x) d\varphi_l^*(x) = \sum_{k = \lceil \operatorname{Log} a \rceil + 1}^{\lceil \operatorname{Log} b \rceil} \frac{W(2^{k-l})}{2^k}.$$

So, by Lemma 4(II),

$$V(\psi_a^*, \varphi^*, \psi_b^*) = V(a_3) + \frac{W(a)}{2^{\lceil \operatorname{Log} a \rceil}} + \sum_{k = \lceil \operatorname{Log} a \rceil + 1}^{\lceil \operatorname{Log} b \rceil} \frac{W(2^{k-l})}{2^k} - V(b_3) - \frac{W(b)}{2^{\lceil \operatorname{Log} b \rceil}} =$$

$$= V(a_3) + \sum_{k = \lceil \operatorname{Log} a \rceil}^{\lceil \operatorname{Log} b \rceil - 1} \frac{W(2^{k-l})}{2^k} - V(b_3),$$

which is $V(l, a_3, b_3)$ since $\lambda(a_3) = 0 = \lambda(b_3)$.

Finally, let $a_4^{0-} \sqcap b_4^{0-} \in \mathcal{C}_6$ or $a_4^{0-} \sqcap b_4^{0-} \in \mathcal{C}_{19}$ be satisfied. In this case, $\gamma_a = 0 = \gamma_b$ and $l = 1$ not only in $\mathcal{C}_6$, but also in $\mathcal{C}_{19}$, and so $m_c' = \lfloor \gamma_c - z_{n'}(c) \rfloor = -2$ for all large n', $c = a, b$. Thus, $\operatorname{Log} a, \operatorname{Log} b \in \mathbf{Z}$ and $\lceil \operatorname{Log} b \rceil = \operatorname{Log} b = r + \operatorname{Log} a = r + \lceil \operatorname{Log} a \rceil$, where $r = 1$ for $\mathcal{C}_6$ and $r = 2$ for $\mathcal{C}_{19}$, and

$\delta_l(a) = 0$, $\delta_l(b) = 2^{-\lceil \operatorname{Log} b \rceil - 2}$ and

$$2 \int_{[a,b]} W(x) d\varphi_l^*(x) = \sum_{k=\lceil \operatorname{Log} a \rceil + 1}^{\lceil \operatorname{Log} b \rceil} \frac{W(2^k)}{2^{k+1}}.$$

So, by Lemma 4(II) again,

$$V(\psi_a^*, \varphi^*, \psi_b^*) = V(a_4) + \frac{W(a)}{2^{\lceil \operatorname{Log} a \rceil + 1}} + \sum_{k=\lceil \operatorname{Log} a \rceil + 1}^{\lceil \operatorname{Log} b \rceil} \frac{W(2^k)}{2^{k+1}} - V(b_4) - \frac{W(b)}{2^{\lceil \operatorname{Log} b \rceil + 1}} =$$

$$= V(a_4) + \sum_{k=\lceil \operatorname{Log} a \rceil}^{\lceil \operatorname{Log} b \rceil - 1} \frac{W(2^k)}{2^{k+1}} - V(b_4) =$$

$$= V(a_4) + \sum_{k=\lceil \operatorname{Log} a \rceil + 1}^{\lceil \operatorname{Log} b \rceil} \frac{W(2^{k-1})}{2^k} - V(b_4),$$

which is $V(1, a_4, b_4)$ since $\lambda(a_4) = -1 = \lambda(b_4)$. Therefore, (2.19) is fully established.

(2) Let the subsequence $\{n''\} \subset \{n'\}$ be given. By Lemma 4(I) and Lemma 3, one can choose a subsequence $\{n'''\} \subset \{n''\}$ such that $\psi_{n'''}^*(c; \cdot) \Longrightarrow \psi_c^*(\cdot)$ as $n''' \to \infty$, for some limiting functions $\psi_c^*(\cdot)$ on $\mathbf{R}$, and $\lfloor \gamma_c - z_{n'''}(c) \rfloor = m_c'''$ for all n''' large enough, where $m_c''' \in \{0, -1, -2\}$ is a constant, $c = a, b$. Then an application of Lemma 2 also gives that $\varphi_{n'''}^*(\cdot) \Longrightarrow \varphi^*(\cdot)$ as $n''' \to \infty$, where $\varphi^*(\cdot)$ is given by (2.14), with m_a''' and m_b''' replacing m_a' and m_b', respectively. If $m = 1, \ldots, 20, 23, 26$, then $\varphi^*(\cdot) \not\equiv 0$ as shown in the proof of Lemma 5. The demonstration below that $(\psi_a^*(\cdot), \varphi^*(\cdot), \psi_b^*(\cdot)) \not\equiv (0,0,0)$ in the remaining cases, that is, when $m = 21, 22, 24, 25, 27, 28$, rests on Lemma 3.

If $\mathbf{C}_{21}$ and $a_2 \sqcap b_2 \in \mathcal{C}_{21}$ hold (along $\{n'''\}$, as will be understood for all conditions of the latter type), then necessarily $(m_a''', m_b''') = (0,0)$, as can be seen from direct inspection of a_2 and b_2, hence $\mathbf{C}_{21}^+$ obtains, and so $\varphi^* \not\equiv 0$ by Lemma 7(2). If $\mathbf{C}_{21}$ and $a_3 \sqcap b_2 \in \mathcal{C}_{21}$ hold, then either $(m_a''', m_b''') = (0,0)$, and so $\varphi^* \not\equiv 0$, or $(m_a''', m_b''') = (-1, 0)$, that is, $\mathbf{C}_{21}^0$ obtains and $\varphi^* \equiv 0$ by Lemma 7(3), but $\psi_a^* \not\equiv 0$ by Lemma 4(2&3). If $\mathbf{C}_{22}$ and $a_1 \sqcap b_1 \in \mathcal{C}_{22}$ hold, then $(m_a''', m_b''') = (-1, -1)$, so we have $\mathbf{C}_{22}^+$ and $\varphi^* \not\equiv 0$ by Lemma 7(2). If $\mathbf{C}_{22}$ and $a_1 \sqcap b_3 \in \mathcal{C}_{22}$ hold, then either $(m_a''', m_b''') = (-1, -1)$, and so we have $\mathbf{C}_{22}^+$ and $\varphi^* \not\equiv 0$ by Lemma 7(2), or $(m_a''', m_b''') = (-1, 0)$, and so $\mathbf{C}_{22}^0$ holds and $\varphi^* \equiv 0$ by Lemma 7(3) but $\psi_b^* \not\equiv 0$ by Lemma 4(2&3). Under $\mathbf{C}_{24}$ and $a_1 \sqcap b_1 \in \mathcal{C}_{24}$, we must have $\mathbf{C}_{24}^+$, and so $\varphi^* \not\equiv 0$. Under $\mathbf{C}_{24}$ and $a_4 \sqcap b_1 \in \mathcal{C}_{24}$, either $\mathbf{C}_{24}^+$ obtains with $\varphi^* \not\equiv 0$, or $\mathbf{C}_{24}^0$ with $\psi_a^* \not\equiv 0$. If we have $\mathbf{C}_{25}$ and $a_2 \sqcap b_1 \in \mathcal{C}_{25}$, then $\mathbf{C}_{25}^+$ obtains with $\varphi^* \not\equiv 0$; while if $\mathbf{C}_{25}$ and $a_2 \sqcap b_3 \in \mathcal{C}_{25}$, then either $\mathbf{C}_{25}^+$

obtains, or $\mathbf{C}_{25}^0$ with $\psi_b^* \not\equiv 0$. Next, $\mathbf{C}_{27}$ and $a_2 \sqcap b_1 \in C_{27}$ give $\mathbf{C}_{27}^+$; while $\mathbf{C}_{27}$ and $a_3 \sqcap b_1 \in C_{27}$ result in either $\mathbf{C}_{27}^+$, or $\mathbf{C}_{27}^0$ with $\psi_a^* \not\equiv 0$. Finally, $\mathbf{C}_{28}$ and the single case $a_1 \sqcap b_4$ in C_{28} yield either $(m_a''', m_b''') = (-1, -2)$, giving $\mathbf{C}_{28}^+$ with $\varphi^* \not\equiv 0$, or $(m_a''', m_b''') = (-1, -1)$, giving $\mathbf{C}_{28}^0$ with $\psi_b^* \not\equiv 0$. □

Proof of Theorem 2. Let $l = \lim_{n' \to \infty} l_{n'}$ and suppose that for some $m = 1, \ldots, 28$ the quadruplet $(\kappa_{a,b}, \gamma_a, \gamma_b, l)$ satisfies condition $\mathbf{C}_m$ and one of the conditions $a_i \sqcap b_j$ in the set C_m along $\{n'\}$. By Lemma 10(2), for every $\{n''\} \subset \subset \{n'\}$ there is an $\{n'''\} \subset \{n''\}$ such that $\varphi_{n'''}^*(\cdot) \Longrightarrow \varphi^*(\cdot)$ and $\psi_{n'''}^*(c; \cdot) \Longrightarrow \Rightarrow \psi_c^*(\cdot)$, $c = a, b$, on $\mathbf{R}$, as $n''' \to \infty$, and $(\psi_a^*(\cdot), \varphi^*(\cdot), \psi_b^*(\cdot)) \not\equiv (0,0,0)$. Then, recalling (1.8), (2.3) and (2.4), so that $\Delta_n^*(a,b)\varphi_n(\cdot)/2^{\lceil \mathrm{Log}(n/k_n) \rceil} = 2\varphi_n^*(\cdot)$ and $\Delta_n^*(a,b)\psi_n(c;\cdot)/2^{\lceil \mathrm{Log}(n/k_n) \rceil} = 2\psi_n^*(c;\cdot)/2^{\lceil \mathrm{Log}\, c \rceil}$, $c = a, b$, where by (2.2) and (2.6), $\inf\{\Delta_n^*(a,b)/2^{\lceil \mathrm{Log}(n/k_n) \rceil} : n \in \mathbf{N}\} > 0$ and $\sup\{\Delta_n^*(a,b)/2^{\lceil \mathrm{Log}(n/k_n) \rceil} : n \in \mathbf{N}\} < \infty$, Theorem 1(i) in [2] and the statement (2.19) in Lemma 10(2) yield $W_{n'''} \xrightarrow{\mathcal{D}} V(\psi_a^*, \varphi^*, \psi_b^*) = V(l, a_i, b_j)$ as $n''' \to \infty$. Since the limit is the same for all such subsequences, we must have $W_{n'} \xrightarrow{\mathcal{D}} V(l, a_i, b_j)$ as $n' \to \infty$.

Conversely, if $W_{n'} \xrightarrow{\mathcal{D}} V$ as $n' \to \infty$ for a non-degenerate V and $\{n''\} \subset \subset \{n'\}$, then, by Theorem 2 in [2], one can find an $\{n'''\} \subset \{n''\}$ such that $\varphi_{n'''}^*(\cdot) \Longrightarrow \varphi^*(\cdot)$ and $\psi_{n'''}^*(c;\cdot) \Longrightarrow \psi_c^*(\cdot)$, $c = a, b$, on $\mathbf{R}$, as $n''' \to \infty$, and $(\psi_a^*(\cdot), \varphi^*(\cdot), \psi_b^*(\cdot)) \not\equiv (0,0,0)$. By Lemma 10(1), $l''' = \lim_{n''' \to \infty} l_{n'''}$ also exists and for some $m_{'''} = 1, \ldots, 28$ the quadruplet $(\kappa_{a,b}, \gamma_a, \gamma_b, l''')$ satisfies condition $\mathbf{C}_{m_{'''}}$ and one of the conditions $a_i''' \sqcap b_j'''$ in the set $C_{m_{'''}}$ along $\{n'''\}$, and, by the already established direct half of the theorem, $V \overset{\mathcal{D}}{=} V(l''', a_i''', b_j''')$. The last statement of the theorem follows from this. □

REFERENCES

[1] Csörgő, S. and Dodunekova, R., Limit theorems for the Petersburg game, *Sums, trimmed sums and extremes*, Progress in Probability, Vol. 23, Birkhäuser, Boston, 1991, 285–315. *MR* **92h**:60031

[2] Csörgő, S. and Mason, D. M., The asymptotic distribution of intermediate sums, *Ann. Probab.* **22** (1994), 145–159.

[3] Csörgő, S., Haeusler, E. and Mason, D. M., The quantile-transform — empirical-process approach to limit theorems for sums of order statistics, *Sums, trimmed sums and extremes*, Progress in Probability, Vol. 23, Birkhäuser, Boston, 1991, 215–267. *MR* **92i**:62101

[4] Feller, W., *An introduction to probability theory and its applications*, I, Wiley, New York, 1950. [Second and third editions: 1957 and 1968; Chapter X, Section 4, in all editions.] *MR* **12**-424; **19**-466; **37** # 3604

[5] Martin-Löf, A., A limit theorem which clarifies the "Petersburg paradox", *J. Appl. Probab.* **22** (1985), 634–643. *MR* **86j**:60074

[6] Gnedenko, B. V. and Kolmogorov, A. N., *Limit distributions for sums of independent random variables*, Addison-Wesley, Cambridge, Mass., 1954. [Revised edition: 1968.] *MR* **16**-52; **38** # 1722

(Received June 10, 1994)

DEPARTMENT OF STATISTICS
THE UNIVERSITY OF MICHIGAN
419 SOUTH STATE STREET
1440 MASON HALL
ANN ARBOR, MI 48109–1027
U.S.A.

DEPARTMENT OF MATHEMATICS
CHALMERS UNIVERSITY OF TECHNOLOGY AND
THE UNIVERSITY OF GÖTEBORG
S-412 96 GÖTEBORG
SWEDEN

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 121–144

STRONG LAWS FOR EXTREME VALUES OF SEQUENCES OF PARTIAL SUM PROCESSES

P. DEHEUVELS and J. STEINEBACH

Dedicated to Pál Révész for his sixtieth birthday

Abstract

Let $\{S_{i,n} : i = 1, 2, \dots\}$ be i.i.d. partial sum processes, each based on the first n observations from a sequence of i.i.d. centered and unit variance random variables. We investigate the limiting behaviour of $M_n(k_n) = \max_{1 \leqq i \leqq k_n} S_{i,n}$, where $\{k_n : n \geqq 1\}$ is a sequence of constants tending to infinity. Under appropriate moment and regularity conditions, we show that, when $n^{-1} \log k_n \to 0$, $(2n \log k_n)^{-1/2} M_n(k_n) \to 1$ a.s., whereas, when $n^{-1} \log k_n \to c \in (0, \infty)$, $(2n \log k_n)^{-1/2} M_n(k_n) \to (2c)^{-1/2} \gamma(c)$ a.s.. Here, γ is a function which characterizes the summand distribution. In the remaining case where $n^{-1} \log k_n \to \infty$, the first-order limiting behaviour of $M_n(k_n)$ is shown to bring only limited information on the underlying distributions.

1. Introduction and statement of main results

Let $\{X_{i,j} : i \geqq 1, j \geqq 1\}$ denote an array of independent and identically distributed [i.i.d.] random variables. For each integer $n \geqq 1$, $i \geqq 1$ and real $k \geqq 1$, set $S_{i,n} = \sum_{j=1}^{n} X_{i,j}$, and $M_n(k) = \max_{1 \leqq i \leqq k} S_{i,n}$. In this paper we are concerned with the limiting behaviour of $M_n(k_n)$ as $n \to \infty$, where $\{k_n : n \geqq 1\}$ is a sequence of real numbers which will be throughout and unless otherwise specified supposed to be nondecreasing and greater than or equal to one.

A simple example motivating the study of statistics such as $M_n(k)$ is to be found in *computer science*, when one seeks to plot a function $H(x)$ at k selected points $x_1, \dots, x_k$. We suppose that, for each x, a computer code may be used to provide an estimation $\hat{H}(x)$ of the exact value $H(x)$. We assume further that the computational steps required to evaluate $\hat{H}(x)$ generate a sequence of n "small errors" which add up to form $\hat{H}(x) - H(x)$. The problem to be addressed is to find the order of magnitude of the *maximal plotting error* $\mathcal{E} := \max_{1 \leqq i \leqq k} (\hat{H}(x_i) - H(x_i))$. If we assume that the "small errors" constitute an array of i.i.d. random variables with the distribution of $X := X_{1,1}$, we

1991 *Mathematics Subject Classification*. Primary 60F15; Secondary 60G15.

Key words and phrases. Law of the iterated logarithm, double arrays, almost sure convergence, strong laws.

0081–6906/96/\$ 5.00 ©1996 *Akadémiai Kiadó, Budapest*

see that the evaluation of $\mathcal{E}$ reduces to that of $M_n(k)$. A discussion of how our results may be used to answer the questions raised by this application is postponed until the end of this section.

Investigations related to this model have been made in a series of publications which we now review. The case where X follows a normal $N(0,1)$ distribution plays here a central role. Under this assumption, we may set without loss of generality $S_{i,n} = W_i(n)$ for $i \geqq 1$ and $n \geqq 1$, where $W_1, W_2, \ldots$ is an i.i.d. sequence of standard Wiener processes on $[0,\infty)$. The study of $M_n(k)$ then becomes a part of a general description of the random sets of functions of $t \in [0,1]$ defined by $\mathbb{E}_n(k) = \{n^{-1/2}W_i(nt) \colon 1 \leqq i \leqq k\}$ for $n \geqq 1$ and $k \geqq 1$. Le Page and Schreiber [15], [16], [17] have considered this question and given to $\mathbb{E}_n(k)$ the physical image of *rocket plumes* (see also de Acosta and Kuelbs [1]). Their results were extended by Deheuvels and Révész [8], who proved Theorem 1.1 below.

For the statement of this theorem, we need some notation. Let $\mathbb{B}[0,1]$ be the set of bounded functions on $[0,1]$, endowed with the topology defined by $\|f\| = \sup_{0\leqq t\leqq 1} |f(t)|$. For $A \subseteqq \mathbb{B}[0,1]$ and $0 < \varepsilon \leqq \infty$, set $A^\varepsilon = \{g \in \mathbb{B}[0,1] \colon \|f-g\| < \varepsilon$ for some $g \in A\}$, and define the corresponding Hausdorff set-metric by setting, for any $A, B \subseteqq \mathbb{B}[0,1]$, $\Delta(A,B) = \inf\{\varepsilon > 0 \colon A \subseteqq B^\varepsilon, B \subseteqq A^\varepsilon\}$. Let J denote the functional defined on $\mathbb{B}[0,1]$ by $J(f) = \int_0^1 \dot{f}^2(t)dt$ when $f \in \mathbb{B}[0,1]$ is absolutely continuous on $[0,1]$ with Lebesgue derivative $\dot{f}$, and $J(f) = \infty$ else, and set $\mathbb{S} = \{f \in \mathbb{B}[0,1] \colon f(0) = 0$ and $J(f) \leqq 1\}$. Finally, denote by $\log_j$ the j-th iterated logarithm.

THEOREM 1.1. *Let* $\{k_n : n \geqq 1\}$ *be such that*

$$(1.1) \qquad \text{(i)} \quad (\log k_n)/\log_2 n \to \infty \quad \textit{as} \quad n \to \infty; \quad \text{(ii)} \quad n^{-1}\log k_n \downarrow .$$

Then, we have

$$(1.2) \qquad \lim_{n\to\infty} \Delta\left((2\log k_n)^{-1/2}\mathbb{E}_n(k_n), \mathbb{S}\right) = 0 \quad \textit{a.s.}.$$

PROOF. See Theorem 1.2 of Deheuvels and Révész [8]. □

Let Θ be a continuous real-valued functional defined on $\mathbb{B}(0,1)$. A simple argument shows that, whenever (1.2) holds, we have

$$(1.3) \qquad \lim_{n\to\infty} (2n\log k_n)^{-1/2} \sup_{f\in\mathbb{E}_n(k_n)} \Theta(f) = \sup_{f\in\mathbb{S}} \Theta(f) \quad \text{a.s.}.$$

By choosing $\Theta(f) = \sup_{0\leqq x\leqq 1} f(x)$ in (1.3), we obtain as a corollary of Theorem 1.1 that, whenever X is $N(0,1)$ and the sequence $\{k_n : n \geqq 1\}$ satisfies (1.1) we have

$$(1.4) \qquad \lim_{n\to\infty} (2n\log k_n)^{-1/2} M_n(k_n) = 1 \quad \text{a.s.}.$$

In view of the well-known invariance principles obtained via approximations of partial sums by Wiener processes (see e.g. [9], [10], [14], [18]), one may expect versions of (1.4) to hold when X follows a general distribution satisfying

$$\mathbf{E}X=0 \quad \text{and} \quad \mathbf{E}X^2=1, \tag{1.5}$$

together with some appropriate additional conditions. Deheuvels and Teicher [7] have followed this line of investigation and given a quite thorough description of the limiting behaviour of $M_n(k_n)$ for *small sequences* such that $(\log k_n)/\log_2 n \to d < \infty$. They also obtained the following partial result concerning the case where $(\log k_n)/\log_2 n \to \infty$.

THEOREM 1.2. *Let $k_n = n^\alpha$ for some $0<\alpha<\infty$. Then, if $\mathbf{E}X=0$, $\mathbf{E}X^2=1$ and $\mathbf{E}(|X|^{2+2\alpha+\varepsilon})<\infty$ for some $\varepsilon>0$, we have* (1.4), *i.e.*

$$\lim_{n\to\infty}(2n\log k_n)^{-1/2}M_n(k_n)=1 \quad a.s..$$

PROOF. See Corollary 5.1 of Deheuvels and Teicher [7], and Theorem 1.6 in the sequel. □

When X is $N(0,1)$ and $k_n=n^\alpha$ for some $0<\alpha<\infty$, (1.1) holds and Theorem 1.1 implies (1.4). Theorem 1.2 extends the validity of this conclusion to the case where X satisfies (1.5) together with $\mathbf{E}|X|^{2+2\alpha+\varepsilon}<\infty$ for some $\varepsilon>0$. As follows from Lemma 2.3 of [7], the latter condition is sharp in the sense that (1.4) fails to hold in general for distributions such that $\mathbf{E}|X|^{2+2\alpha-\varepsilon}=\infty$ for some $\varepsilon>0$. One of the purposes of this paper is to make precise this statement. This will be achieved in the forthcoming Theorems 1.6 and 2.3. In the special case where $\alpha=1$ (that is for $k_n=n$) results of this kind have been obtained by Cramér [4] and Esséen [12].

Besides the purpose of describing the behaviour of $M_n(k_n)$ in the range corresponding to $k_n=n^\alpha$, our main line of investigation is motivated by the following observation. The general meaning of Theorem 1.2 is that, as long as X has finite moments of sufficient order, the limiting first-order asymptotic behaviour of $M_n(n^\alpha)$ depends upon the distribution of X *only through the values,* $\mathbf{E}X=0$ *and* $\mathbf{E}X^2=1$, *of the first two moments.* It is therefore natural to ask whether there exist sequences $\{k_n : n \geqq 1\}$ satisfying (1.1), but for which this *invariance property* fails.

Our first result, stated in Theorem 1.3 below, gives a positive answer to this question. For the statement of this theorem, the following additional notation is needed. Letting $\varphi(t)=\mathbf{E}e^{tX}$ denote the moment generating function [m.g.f.] of X, introduce the assumptions:

(H1) $\mathbf{E}X=0$;

(H2) The distribution of X is nondegenerate, i.e. $\mathbf{P}(X=x)<1$ for all x;

(H3) $t_1:=\inf\{t: \varphi(t)<\infty\}\leqq 0<t_0:=\sup\{t: \varphi(t)<\infty\}$.

The *Chernoff function* ψ of X and its inverse γ on $[0,\infty)$, are given respectively by

$$(1.6) \qquad \psi(\alpha) = \sup_{t:\ \phi(t)<\infty} \{t\alpha - \log\phi(t)\} \quad \text{for} \quad -\infty < \alpha < \infty,$$

and

$$(1.7) \qquad \gamma(c) = \sup\{\alpha \geqq 0 : \psi(\alpha) \leqq c\} \quad \text{for} \quad c \geqq 0.$$

Denote by $\omega = \text{ess sup}\, X = \sup\{x\colon F(x) < 1\}$, with $F(x) = \mathbf{P}(X \leqq x)$, the (eventually infinite) upper endpoint of the distribution of X. The assumptions (H1)–(H2) jointly imply that $0 < \omega \leqq \infty$. Moreover, it follows from Lemmas 2.1 and 2.2 of Deheuvels [5] that, under (H1)–(H2)–(H3), ψ is finite, increasing and strictly convex on $[0,\omega)$, infinite on (ω,∞), and nondecreasing and convex on $[0,\infty)$. In addition, under the sole assumption (H1), ψ is nonnegative and convex on $(-\infty,\infty)$, with

$$(1.8) \qquad \text{(i) } \psi(0) = 0; \quad \text{(ii) } \lim_{\alpha\to\infty} \alpha^{-1}\psi(\alpha) = t_0; \quad \text{(iii) } \lim_{\alpha\to-\infty} \alpha^{-1}\psi(\alpha) = t_1.$$

When (H1) holds but not (H2), $X = 0$ a.s., so that $\psi(\alpha) = 0$ for $\alpha = 0$ and $\psi(\alpha) = \infty$ else. In this case, $\gamma(c) = 0$ for all $c \geqq 0$. When (H1)–(H2) hold, but not (H3) (with $t_1 = 0$), $t_0 = 0$ and $\psi(\alpha) = 0$ for all $\alpha \in \mathbb{R}$. In this case, $\gamma(c) = 0$ for $c = 0$ and $\gamma(c) = \infty$ else. When (H1)–(H2)–(H3) hold, it is readily inferred from (1.7) and the properties of ψ that γ is a concave function on $[0,\infty)$ such that, for all $c > 0$,

$$(1.9) \qquad \text{(i) } \gamma(0) = 0 < \gamma(c) < \infty; \qquad \text{(ii) } 0 < \gamma(c) \leqq \omega = \lim_{z\to\infty} \gamma(z),$$

and, for all $0 < \varepsilon < \gamma(c)$,

$$(1.10) \qquad 0 < \psi(\gamma(c) - \varepsilon) < c < \psi(\gamma(c) + \varepsilon) \leqq \infty.$$

We may now state our main result corresponding to *large sequences*, typically of the form $k_n = e^{nc}$ for some $c > 0$.

THEOREM 1.3. *Assume (H1)–(H2)–(H3), and let* $\{k_n\colon n \geqq 1\}$ *be such that*

$$(1.11) \qquad n^{-1}\log k_n \to c \in (0,\infty).$$

Then, we have

$$(1.12) \qquad \lim_{n\to\infty} (2n\log k_n)^{-1/2} M_n(k_n) = (2c)^{-1/2}\gamma(c) \quad a.s..$$

In Section 2, we will provide a proof of Theorem 1.3, together with a discussion of the sharpness of the assumptions which are made in this statement. In particular, we will show that, in general, whenever (1.11) and (H1)–(H2) hold but not (H3), the sequence $(2n \log k_n)^{-1/2} M_n(k_n)$ is unbounded with probability one (see Remark 2.2 in the sequel).

Theorem 1.3 shows that, under (H1)–(H2)–(H3) and for *large sequences* of the form $k_n = e^{nc}$ with $c > 0$, the first-order limiting behaviour of $M_n(k_n)$ is dependent of characteristics of the distribution of X *other* than the first two moments. This situation is the opposite of that, previously described, which is covered by the assumptions of Theorem 1.2. In view of the results of Bártfai [2], showing that $\{\gamma(c) : c > 0\}$ completely determines $F(x) = \mathbf{P}(X \leqq x)$, we obtain via (1.12) that *the knowledge of the limit of* $n^{-1} M_n(e^{nc})$ *for each* $c > 0$ *specifies* F. This phenomenon is very similar to that which is described by the Erdős–Rényi strong law of large numbers for maximal increments of partial sums. The latter strong law gave rise to a huge literature (see e.g. Erdős and Rényi [11], Deheuvels [5] and the references therein). Interestingly, when X is $N(0,1)$, we have $(2c)^{-1/2}\gamma(c) = 1$ for all $c > 0$, so that (1.12) then coincides with (1.4). In this particular case, and since the sequence $k_n = e^{nc}$ satisfies (1.1), we may therefore also infer (1.12) from Theorem 1.1.

A natural question raised by the conclusion of Theorem 1.3 is to ask whether the *distribution-dependent* properties of $M_n(k_n)$ which hold for sequences $\{k_n : n \geqq 1\}$ such that $n^{-1} \log k_n \to c \in (0, \infty)$ are shared by other families of sequences with either a *larger* or a *smaller* order of magnitude. We will now show that it is essentially not the case. We will treat successively the *very large* sequences for which $n^{-1} \log k_n \to \infty$, and then, the case of *moderate to small* sequences for which $n^{-1} \log k_n \to 0$.

Our next theorem corresponds to *very large sequences* characterized by the assumption that $n^{-1} \log k_n \to \infty$. Recall from our discussion preceding Theorem 1.3 the notation $\omega = \operatorname{ess\,sup} X = \sup\{x : F(x) < 1\}$ with $F(x) = \mathbf{P}(X \leqq x)$, and the observation that, under (H1), the assumption (H2) that the distribution of X is nondegenerate is equivalent to the condition $\omega > 0$.

THEOREM 1.4. *Assume (H1)–(H2)–(H3), and let* $\{k_n : n \geqq 1\}$ *be such that* $n^{-1} \log k_n \downarrow$ *and*

$$n^{-1} \log k_n \to \infty. \tag{1.13}$$

(1) *If* $\omega = \operatorname{ess\,sup} X < \infty$, *then*

$$\lim_{n\to\infty} \left\{ \frac{M_n(k_n)}{n\omega} \right\} = \lim_{n\to\infty} \left\{ \frac{M_n(k_n)}{n\gamma(n^{-1}\log k_n)} \right\} = 1 \quad a.s.. \tag{1.14}$$

(2) *If* $\omega = \operatorname{ess\,sup} X = \infty$ *and either* $(\log k_{n+1})/\log k_n \to 1$ *as* $n \to \infty$, *or*

$$(-\log(1 - F(x)))/\psi(x) \to 1 \quad as \quad x \to \infty, \tag{H4}$$

then

$$\limsup_{n\to\infty} \left\{ \frac{M_n(k_n)}{n\gamma(n^{-1}\log k_n)} \right\} = 1 \quad a.s.. \tag{1.15}$$

(3) *If* $\omega = \text{ess sup } X = \infty$ *and if*

$$\text{(H5)} \qquad \lim_{x\to\infty} x^{-1}\gamma(-\log(1-F(x))) = 1$$

then (1.15) *holds with* lim sup *replaced by* lim.

In Section 2 the proof of Theorem 1.4 will be given together with other related results of interest. We note (see e.g. Mason [19]) that the assumption (H4) implies (H5), whereas the converse is not true (see e.g. Remark 2.4). In Case (1) of Theorem 1.4, we see that the first-order limiting behaviour of $M_n(k_n)$ depends then only upon ω and brings therefore only *limited information upon F*. A similar observation can be made in Case (2). It is noteworthy that these results are analogues of that obtained by Mason [19] for large increments of partial sums.

Our last series of theorems corresponds to *moderate to small sequences*, characterized by the assumption that $n^{-1}\log k_n \to 0$ as $n \to 0$. For the classes of sequences we consider, we will not only show that the first-order limiting behaviour of $M_n(k_n)$ depends upon the distribution of X only through $\mathbf{E}X = 0$ and $\mathbf{E}X^2 = 1$, but also derive sharp moment-type conditions to ensure the validity of (1.3). Our main results concerning this range are captured in the following two theorems, the second of which brings a refinement to Theorem B. Let $a\wedge b = \min(a,b)$, $a \vee b = \max(a,b)$ and $\log_+ u = \log(u \vee e)$.

THEOREM 1.5. *Let $\{k_n : n \geqq 1\}$ be such that, for some $\beta \in (0,1)$,*

$$\text{(1.16)} \qquad n^{-\beta}\log k_n \to \rho \in (0,\infty) \quad as \quad n\to\infty.$$

Assume further that

$$\text{(H6)} \qquad \mathbf{E}\{\exp(\lambda |X|^{\frac{2\beta}{\beta+1}})\} < \infty \quad for\ some \quad \lambda > 0.$$

Then (1.4) *holds, i.e.*

$$\lim_{n\to\infty} (2n\log k_n)^{-1/2} M_n(k_n) = 1 \quad a.s..$$

THEOREM 1.6. *Let $\{k_n : n \geqq 1\}$ be such that, for some $\alpha > 0$,*

$$\text{(1.17)} \qquad 0 < \liminf \left\{\frac{k_n}{n^\alpha}\right\} \leqq \limsup \left\{\frac{k_n}{n^\alpha}\right\} < \infty.$$

Assume further that

$$\text{(H7)} \qquad \mathbf{E}\{|X|^{2\alpha+2}/(\log_+ |X|)^{\alpha+1}\} < \infty.$$

Then (1.4) *holds, i.e.*

$$\lim_{n\to\infty} (2n\log k_n)^{-1/2} M_n(k_n) = 1 \quad a.s..$$

The proof of Theorems 1.5 and 1.6 will be given in the forthcoming Section 2. In this section, we will also show that the assumptions of these theorems are sharp in the sense that (1.4) fails to hold in general in each of the following two cases: (i) (1.16) holds but not (H6); (ii) (1.17) holds but not (H7).

It will become obvious from the arguments of our proofs in the sequel that our methods enable to describe as well the limiting behaviour of $M_n(k_n)$ for other families of sequences than those we consider here, corresponding to the assumptions (1.11), (1.13), (1.16) and (1.17). Such extensions are left to the reader and we will limit ourselves to the study of the most important examples. Our main purpose will be to describe in each case the information carried by $M_n(k_n)$ on the distribution of X.

We need here mention a connection of the preceding results with the *large deviation principle* [LDP] obtained recently by Deheuvels and Vervaat (1994). For this sake, we return to the setting of Theorem B, and consider the functions of $t \in [0,1]$ defined for $n \geqq 1$ and $i \geqq 1$ by

$$Y_{n,i}(t) = n^{-1/2} W_i(nt). \tag{1.18}$$

It is well-known that, for each fixed $i \geqq 1$, $Y_{n,i}$ satisfies an LDP, the most classical form of which is due to Schilder [21] (see e.g. Fact 2 in Deheuvels and Révész [8]). The latter LDP may be formulated in concise notation (see e.g. Deheuvels and Vervaat [6]) as follows. For each $i \geqq 1$, we have, as $n \to \infty$,

$$\text{law}^{1/n} Y_{n,i} \to e^{-J} \quad \text{narrowly}. \tag{1.19}$$

Whenever such a statement holds, Deheuvels and Vervaat [6] have shown that there exists a subsequent LDP for the random measure

$$\sum_{1 \leqq i \leqq \exp(n^c)} \delta_{Y_{n,i}}$$

where δ_f denotes here the Dirac measure concentrated at f. By making use of such LDP's, it is possible to derive a new proof of Theorem 1.1. Moreover, functional versions of Theorems 1.3 and 1.4 could be obtained along the same lines, given the appropriate LDP's for the component partial sum processes. However, such LDP's are available only in special cases (see e.g. Deheuvels [5], Mogulskii [20] and the references therein), which do not cover yet the general case of partial sum processes satisfying (H1)–(H2)–(H3). Therefore, such an approach does not lead at present to alternate proofs of our results.

To conclude this section, we return to the numerical analysis example which was introduced at the beginning, and consider the case where the distribution of an individual "small error" X has a finite moment-generating function $\varphi(s)$ in a neighbourhood of 0. In most cases of interest, $\log n$ is of the same order of magnitude as $\log k$, so that we may consider ourselves in the range where $(\log k)/\log n$ is bounded away from 0 and infinity. Under these assumptions, it is easily inferred from our theorems that the thumb-rule evaluation of $\mathcal{E}$ given by

$$\mathcal{E} \approx \sigma\sqrt{2n\log k}, \tag{1.20}$$

where $\sigma^2 = \operatorname{Var} X$, is asymptotically valid when both n and k become large. It does, however, become incorrect when $\log n$ is small relative to $\log k$, in which case one should prefer the alternative evaluation

$$\mathcal{E} \approx n\gamma(n^{-1}\log k). \tag{1.21}$$

Making use of the easily verified fact that $\gamma(c) = (1+o(1))\sigma\sqrt{2c}$ as $c \to 0$, we see that (1.20) and (1.21) are in agreement when $(\log k)/\log n$ is bounded away from 0 and infinity. For example, when X is uniformly distributed on $\left[-\frac{\delta}{2}, \frac{\delta}{2}\right]$, with δ denoting the precision of the computer, we get $\sigma = \delta/\sqrt{12}$ and get the approximation, based on (1.20),

$$\mathcal{E} \approx \delta\sqrt{\frac{1}{6}n\log k}. \tag{1.22}$$

To illustrate (1.22), let $\delta = 2^{-32} \approx 2.33 \times 10^{-10}$, and choose $k = 100$, $n = 1000$. We obtain immediately via this formula, $\mathcal{E} \approx 27.7 \times \delta \approx 6.45 \times 10^{-9}$.

In Section 2 below, we give proofs of Theorems 1.3–1.6, together with a discussion of the sharpness of the assumptions made in these theorems, and some additional related results of interest.

2. Proofs and additional results

2.1. Large sequences. We inherit the notation of Section 1, and assume in this subsection, throughout and unless otherwise specified, that (H1)–(H2)–(H3) hold. We consider here the case of *large sequences*, characterized by the assumption that, as $n \to \infty$,

$$n^{-1}\log k_n \to c \in (0,\infty). \tag{2.1}$$

Throughout, we will suppose that $k_n \geqq 1$ for all $n \geqq 1$. On the other hand, we will not need the condition that k_n is nondecreasing and let therefore this sequence be *eventually non-monotonic*. We will start by proving Theorem 1.3, and then establish the sharpness of the assumptions of this theorem (see e.g. Theorem 2.1 in the sequel).

The following three lemmas are directed towards the proof of Theorem 1.3. We start by giving classical large deviation bounds which are instrumental for our needs.

LEMMA 2.1. *Under (H1)–(H2)–(H3), for any $\alpha \geqq 0$ and $\eta > 0$, we have for all n sufficiently large the inequalities*

$$\exp\left(-n\{\psi(\alpha)+\eta\}\right) \leqq \mathbf{P}(S_{1,n} \geqq n\alpha) \leqq \exp\left(-n\{\psi(\alpha)\}\right). \tag{2.2}$$

PROOF. Both inequalities are trivial when $\alpha = 0$. When $\alpha > 0$, (2.2) is a consequence of Theorem 1 of Chernoff [3]. □

LEMMA 2.2. *Under (H1), (H2), and* (2.1), *we have*

$$\limsup_{n\to\infty} n^{-1} M_n(k_n) \leqq \gamma(c) \quad a.s.. \tag{2.3}$$

PROOF. When (H1)–(H2) hold but not (H3), we have $\gamma(c) = \infty$ and (2.3) is trivial. We may therefore assume without loss of generality that (H1)–(H2)–(H3) hold. Fix an arbitrary $\varepsilon > 0$. It follows from (1.10) that $c < \psi(\gamma(c)+\varepsilon) \leqq \infty$. We now choose $\theta = \frac{1}{2}\{\psi(\gamma(c)+\varepsilon) - c\}$ when $\psi(\gamma(c)+\varepsilon) < < \infty$, and set $\theta = 1$ otherwise. In both cases, $0 < \theta < \infty$, and $\psi(\gamma(c)+\varepsilon) > c + \theta > c$. It follows from these inequalities and (2.1) that there exists an n_0 such that, for all $n \geqq n_0$, $\log k_n < n(c + \frac{\theta}{2})$, and hence, $\log k_n - n\psi(\gamma(c)+\varepsilon) < -n\frac{\theta}{2}$. By combining this last inequality with (2.2) taken with $\alpha = \gamma(c) + \varepsilon$, we obtain readily that, for all large n,

$$\begin{aligned}\mathbf{P}(M_n(k_n) \geqq n(\gamma(c)+\varepsilon)) &\leqq k_n \mathbf{P}(S_{1,n} \geqq n(\gamma(c)+\varepsilon)) \\ &\leqq \exp\left(\log k_n - n\psi(\gamma(c)+\varepsilon)\right) < \exp\left(-n\frac{\theta}{2}\right).\end{aligned}$$

Thus, since $\sum_{n=1}^{\infty} \exp\left(-n\frac{\theta}{2}\right) < \infty$, the Borel–Cantelli lemma, when applied to the events $\{M_n(k_n) \geqq n(\gamma(c)+\varepsilon)\}$, shows that the left-hand side of (2.3) is almost surely less than or equal to $\gamma(c) + \varepsilon$. The proof is completed by the observation that $\varepsilon > 0$ may be chosen arbitrarily small. □

REMARK 2.1. The conclusion (2.3) of Lemma 2.2 remains valid if we only assume that (i) The random variables $\{X_{i,j} \colon i \geqq 1,\, j \geqq 1\}$ are identically distributed; (ii) For each $i \geqq 1$, the random variables $\{X_{i,j} \colon j \geqq 1\}$ are independent.

LEMMA 2.3. *Under (H1), (H3) and* (2.1), *we have*

$$\liminf_{n\to\infty} n^{-1} M_n(k_n) \geqq \gamma(c) \quad a.s.. \tag{2.4}$$

PROOF. If X is degenerate, then (H1) implies that $X = 0$ a.s., so that both sides of (2.4) are equal to 0 and the lemma is trivial. We may assume therefore, without loss of generality, that (H1)–(H2)–(H3) hold. Recalling from (1.9)(i) that, in this case, $0 < \gamma(c) < \infty$, fix an arbitrary $\varepsilon \in (0, \gamma(c))$. By (1.10), we have $0 < \psi(\gamma(c) - \varepsilon) =: c - 2\eta < c$. This last inequality, in combination with (2.2), taken with $\alpha = \gamma(c) - \varepsilon$, and the fact that $\{1-x\}^r \leqq \exp(-rx)$ for all $r > 0$ and $0 \leqq x \leqq 1$, shows that for all large n

$$\begin{aligned} \mathbf{P}(M_n(k_n) < n(\gamma(c)-\varepsilon)) &= \{1 - \mathbf{P}(S_{1,n} \geqq n(\gamma(c)-\varepsilon))\}^{\lfloor k_n \rfloor} \\ &\leqq \exp(-(k_n - 1)\mathbf{P}(S_{1,n} \geqq n(\gamma(c)-\varepsilon))) \\ &\leqq 3\exp(-k_n \exp(-n(\psi(\gamma(c)-\varepsilon)+\eta))) \\ &\leqq 3\exp(-\exp(\log k_n - n(c-\eta))), \end{aligned} \tag{2.5}$$

where $\lfloor k_n \rfloor > k_n - 1$ denotes the integer part of k_n. Since (2.1) entails that, for all large n, $\log k_n \geqq n(c - \frac{\eta}{2})$, and hence, $\log k_n - n(c-\eta) \geqq n\frac{\eta}{2}$, it follows from (2.5) that, for all n sufficiently large,

$$\mathbf{P}(M_n(k_n) < n(\gamma(c)-\varepsilon)) \leqq 3\exp(-e^{n\frac{\eta}{2}}). \tag{2.6}$$

In view of the fact that $\varepsilon > 0$ in (2.6) may be chosen arbitrarily small, and since $\sum_{n=1}^{\infty} \exp(-e^{n\frac{\eta}{2}}) < \infty$, an application of the Borel–Cantelli lemma completes the proof of (2.4). □

Proof of Theorem 1.3. Combine Lemmas 2.2 and 2.4. □

We now turn to a discussion of the sharpness of the assumptions of Theorem 1.3. Towards this aim, we will establish two technical lemmas, the first of which provides us with a simple but powerful inequality.

Lemma 2.4. *For any $n \geqq 1$, we have*

$$\max_{1 \leqq i \leqq k_n \wedge k_{n+1}} |X_{i,n+1}| \leqq \max_{1 \leqq i \leqq k_{n+1}} |S_{i,n+1}| + \max_{1 \leqq i \leqq k_n} |S_{i,n}|. \tag{2.7}$$

Proof. Since, for $i = 1, 2, \ldots$, and $n = 1, 2, \ldots$, $X_{i,n+1} = S_{i,n+1} - S_{i,n}$, the triangle inequality entails $|X_{i,n+1}| \leqq |S_{i,n+1}| + |S_{i,n}|$. By maximizing both sides of this inequality with respect to i with $1 \leqq i \leqq k_n \wedge k_{n+1}$, we obtain readily (2.7). □

Let $\{m_n : n \geqq 1\}$ denote a sequence of positive constants such that

$$\text{(i)} \quad \limsup_{n\to\infty} (m_{n+1}/m_n) < \infty; \qquad \text{(ii)} \quad m_n \uparrow \infty \text{ as } n \uparrow \infty. \tag{2.8}$$

Lemma 2.5. *Let $\{m_n : n \geqq 1\}$ satisfy (2.8), and be such that*

$$\limsup_{n\to\infty} m_n^{-1} \max_{1 \leqq i \leqq k_n} |S_{i,n}| < \infty \quad a.s.. \tag{2.9}$$

Then, for some $d > 0$, we have

$$\sum_{j=1}^{\infty} \mathbf{P}(dm_j \leqq |X| < dm_{j+1}) \sum_{n=1}^{j} (k_n \wedge k_{n+1}) < \infty. \tag{2.10}$$

PROOF. By (2.7) and (2.8)(i), we see that, whenever (2.9) holds, we have

$$\limsup_{n\to\infty} m_n^{-1} \max_{1\leq i\leq k_n\wedge k_{n+1}} |X_{i,n+1}| < \infty \quad \text{a.s..} \tag{2.11}$$

Since, for each $d \geq 0$, the events $A_n(d) = \max_{1\leq i\leq k_n\wedge k_{n+1}} |X_{i,n+1}| \geq dm_n\}$, $n = 1, 2, \ldots$, are independent, the Borel–Cantelli lemma implies that $\mathbf{P}(A_n(d) \text{ i.o.}) = 0$ or 1 according as $\sum_{n=1}^{\infty} \mathbf{P}(A_n(d)) < \infty$ or $= \infty$. Therefore, (2.11) implies the existence of a value of $d > 0$ for which $\sum_{n=1}^{\infty} \mathbf{P}(A_n(d)) < \infty$. Since then

$$\mathbf{P}(A_n(d)) = 1 - (1 - \mathbf{P}(|X| \geq dm_n))^{\lfloor k_n\wedge k_{n+1}\rfloor} \to 0,$$

we must have $\mathbf{P}(A_n(d)) = (1 + o(1))\lfloor k_n \wedge k_{n+1}\rfloor \mathbf{P}(|X| \geq dm_n)$ as $n \to \infty$. Recalling the assumption that $k_n \geq 1$ for each $n \geq 1$, we see therefore that

$$\sum_{n=1}^{\infty} (k_n \wedge k_{n+1})\mathbf{P}(|X| \geq dm_n) < \infty. \tag{2.12}$$

Now, a straightforward calculus making use of (2.8)(ii) shows that

$$\begin{aligned} &\sum_{n=1}^{\infty} (k_n \wedge k_{n+1})\mathbf{P}(|X| \geq dm_n) = \\ &= \sum_{n=1}^{\infty} (k_n \wedge k_{n+1}) \sum_{j=n}^{\infty} \mathbf{P}(dm_j \leq |X| < dm_{j+1}) = \\ &= \sum_{j=n}^{\infty} \mathbf{P}(dm_j \leq |X| < dm_{j+1}) \sum_{n=1}^{j} (k_n \wedge k_{n+1}). \end{aligned} \tag{2.13}$$

By combining (2.12) and (2.13), we obtain (2.10). □

THEOREM 2.1. *Let $\{k_n : n \geq 1\}$ be such that*

$$n^{-1} \log k_n \to c \in (0, \infty).$$

Then, a necessary and sufficient condition for

$$\limsup_{n\to\infty} (2n \log k_n)^{-1/2} \max_{1\leq i\leq k_n} |S_{i,n}| < \infty \quad a.s. \tag{2.15}$$

is that $\varphi(t) < \infty$ in a neighbourhood of 0.

PROOF. The fact that (2.15) holds when $\varphi(t) < \infty$ in a neighbourhood of 0 being a straightforward consequence of Theorem 1.3, we limit ourselves to the proof of the converse statement. Towards this aim, set $k'_n = \inf\{k_m : m \geqq n\}$, and observe that the sequence $\{k'_n : n \geqq 1\}$ so defined satisfies for each $n \geqq 1$ the inequalities $1 \leqq k'_n \leqq k_n$. Moreover, under (2.1), it is such that $n^{-1} \log k'_n \to c$. In view of the obvious inequality $\max_{1 \leqq i \leqq k'_n} |S_{i,n}| \leqq \max_{1 \leqq i \leqq k_n} |S_{i,n}|$, it follows that there is no loss of generality in assuming in our proof that $k_n = k'_n$, or equivalently, that k_n is nondecreasing.

This being the case, we may choose $m_n = n$ in Lemma 2.5 to obtain , via (2.10) and the inequality $k_j \leqq \sum_{n=1}^{j} k_n$, that, under (2.15) and (2.1) (i.e. when $n^{-1} \log k_n \to c$), we have for some $d > 0$

$$\sum_{j=1}^{\infty} \mathbf{P}(dj \leqq |X| \leqq d(j+1))k_j < \infty. \tag{2.16}$$

Now, (2.1) implies that, for all large $j, k_j \geqq \exp\left(\frac{c}{2d} d(j+1)\right)$, which, when combined with (2.16), implies that $\mathbf{E}\left(\frac{c}{2d}|X|\right) < \infty$, from where the conclusion of the theorem is straightforward. $\square$

REMARK 2.2. Assume that (H1) holds and that $\varphi(t) < \infty$ in a *left* neighbourhood of 0. By an application of Theorem 1.3 with the formal replacement of X by $-X$ we get

$$\limsup_{n \to \infty} n^{-1} \max_{1 \leqq i \leqq k_n} (-S_{i,n}) < \infty \quad \text{a.s.},$$

so that the assumption that

$$\limsup_{n \to \infty} n^{-1} M_n(k_n) < \infty \quad \text{a.s.} \tag{2.18}$$

suffices to imply (2.15). Thus, by Theorem 2.1, $\varphi(t)$ must be also finite in a *right* neighbourhood of 0. Since the converse holds by Theorem 1.3, we see, in this particular case, that the condition (H3) (i.e. $\varphi(t) < \infty$ in a *right* neighbourhood of 0) is *necessary and sufficient* for (2.18). In fact, even without assuming that $\varphi(t) < \infty$ in a left neighbourhood of 0, it is true that, whenever $\varphi(t) = \infty$ for all $t > 0$, we have

$$\limsup_{n \to \infty} n^{-1} M_n(k_n) = \infty \quad \text{a.s.}. \tag{2.19}$$

The proof of this last statement can be obtained along the lines of Steinebach [22], and we omit the details of this argument. This shows that the assumptions of Theorem 1.3 are sharp.

REMARK 2.3. We will not investigate in details the limiting behaviour of $M_n(k_n)$ when $\{k_n : n \geqq 1\}$ satisfies (2.1), and when (H3) does not hold.

Our main result concerning this case will be given in Theorem 2.2 in the sequel and gives only a general asymptotic lower bound for this quantity.

2.2. Very large sequences. In this subsection, we consider *very large sequences*, and assume throughout and unless otherwise specified that

$$n^{-1}\log k_n \to \infty \quad \text{as} \quad n \to \infty. \tag{2.20}$$

The following Lemmas 2.6–2.11 are directed towards proving Theorem 1.4. The technology which we use in their proofs is greatly inspired from that of Mason [19]. However, due to the different setting, some modifications are necessary, and we will give most of the details for the sake of completeness.

Our first lemma, stated below, treats the case where $\omega = \text{ess sup}\, X \in (0,\infty)$.

LEMMA 2.6. *Under (H1)–(H2)–(H3), assume that* (2.20) *holds, and that* $\omega < \infty$. *Then,*

$$\lim_{n\to\infty}\left\{\frac{M_n(k_n)}{n\omega}\right\} = \lim_{n\to\infty}\left\{\frac{M_n(k_n)}{n\gamma(n^{-1}\log k_n)}\right\} = 1 \quad a.s.. \tag{2.21}$$

PROOF. First, notice that, under (H1)–(H2)–(H3), $\omega > 0$. Moreover, by (1.9)(ii), we see that (2.20) entails $\gamma(n^{-1}\log k_n) \to \omega$. Thus, since we always have $n^{-1}M_n(k_n) \leqq \omega$, in order to establish (2.21), we need only prove that

$$\liminf_{n\to\infty}\left\{\frac{M_n(k_n)}{n\omega}\right\} \geqq 1 \quad \text{a.s..} \tag{2.22}$$

Towards this aim, we follow the lines of the proof of (2.1) in [19]. First, we choose an arbitrary $\varepsilon \in (0,\omega)$, and observe that $\delta := \mathbf{P}(X > \omega - \varepsilon) \in (0,1)$. Then, we make use of the inclusion of events $\bigcap_{i=1}^{n}\{X_{1,j} > (\omega-\varepsilon)\} \subseteqq \{S_{1,n} > n(\omega-\varepsilon)\}$, which, after complementation and taking probabilities of both sides, yields the bound, similar to (2.2) of [19],

$$\mathbf{P}(S_{1,n} \leqq n(\omega-\varepsilon)) \leqq 1 - (\mathbf{P}(X > (\omega-\varepsilon))^n = 1 - \delta^n. \tag{2.23}$$

An application of (2.23) shows in turn that

$$\begin{aligned}\mathbf{P}(M_n(k_n) \leqq n(\omega-\varepsilon)) &= \{\mathbf{P}(S_{1,n} \leqq n(\omega-\varepsilon))\}^{k_n} \\ &\leqq \{1-\delta^n\}^{k_n} \leqq \exp(-k_n\delta^n).\end{aligned} \tag{2.24}$$

Observe that, by (2.20), $n^{-1}\log(k_n\delta^n) = n^{-1}\log k_n + \log\delta \to \infty$, so that, for all n sufficiently large, $n^{-1}\log(k_n\delta^n) \geqq \log 2$. This, in turn, suffices to show via (2.24) that

$$\sum_{n=1}^{\infty}\mathbf{P}(M_n(k_n) \leqq n(\omega-\varepsilon)) < \infty.$$

The Borel–Cantelli lemma, in combination with the fact that $\varepsilon > 0$ may be chosen arbitrarily small, completes the proof of (2.22), and hence of (2.21). □

We now turn to the more difficult proof of (1.15), and assume from now on that $\omega = \infty$. We make use of the following lemmas.

LEMMA 2.7. *For any integer* $n \geqq 1$, *real* $m \geqq 0$ *and* $\varepsilon > 0$, *we have*

$$\mathbf{P}\Big(S_{1,n} \geqq (1+\varepsilon)n\gamma\Big(\frac{m}{n}\Big)\Big) \leqq \exp(-(1+\varepsilon)m). \tag{2.25}$$

PROOF. This is Lemma 2.3 of [19], with the formal change of k into n.□

LEMMA 2.8. *Under (H1)–(H2)–(H3), assume that* (2.20) *holds, and that* $\omega = \infty$. *Then,*

$$\limsup_{n\to\infty}\Big\{\frac{M_n(k_n)}{n\gamma(n^{-1}\log k_n)}\Big\} \leqq 1 \quad a.s.. \tag{2.26}$$

PROOF. Fix any $\varepsilon > 0$. Making use of the Bonferroni inequality, and then, of (2.25) taken with $m = \log k_n$, we obtain the chain of inequalities, for $n \geqq 1$ and $k_n \geqq 1$,

$$\begin{aligned}
&\mathbf{P}(M_n(k_n) \geqq (1+\varepsilon)n\gamma(n^{-1}\log k_n)) \\
&\qquad \leqq k_n\mathbf{P}(S_{1,n} \geqq (1+\varepsilon)n\gamma(n^{-1}\log k_n)) \\
&\qquad \leqq k_n\exp(-(1+\varepsilon)\log k_n) = \exp(-\varepsilon\log k_n).
\end{aligned} \tag{2.27}$$

By (2.20), we have $\varepsilon\log k_n \geqq n\log 2$ for all n sufficiently large, whence the right side of (2.27) is summable in n. The Borel–Cantelli lemma in combination with the fact that $\varepsilon > 0$ may be chosen arbitrarily small, completes the proof of (2.26). □

LEMMA 2.9. *Under (H1)–(H2)–(H3), assume that* (2.20) *holds, and that* $\omega = \infty$. *Then,*

$$\liminf_{z\to\infty}\Big\{-\frac{\log(1-F(z))}{\psi(z)}\Big\} = 1. \tag{2.28}$$

PROOF. This is Lemma 2.2 of [19]. Note that the inequality $\mathbf{P}(X \geqq z) \leqq \varphi(t)e^{-tz}$ implies that we always have

$$-\frac{\log(1-F(z))}{\psi(z)} \geqq 1. \tag{2.29}$$
□

LEMMA 2.10. *Under (H1)–(H2)–(H3), assume that* (2.20) *holds, and that* $\omega = \infty$. *Assume further that either one of the conditions* (i) *or* (ii) *below is satisfied.*

(i) $(\log k_{n+1})/\log k_n \to 1$ *as* $n \to \infty$;

(ii) *For each small* $\varepsilon > 0$,

$$\text{(H8)} \qquad \liminf_{n\to\infty}\Big(\inf_{\frac{1}{(1+\varepsilon)^3}\frac{\log k_n}{n}\leqq x\leqq\frac{1}{(1+\varepsilon)^2}\frac{\log k_n}{n}}\Big\{\frac{-\log(1-F(x))}{\psi(x)}\Big\}\Big) = 1.$$

Then,

$$\text{(2.28)} \qquad \limsup_{n\to\infty}\Big\{\frac{M_n(k_n)}{n\gamma(n^{-1}\log k_n)}\Big\} \geqq 1 \quad a.s..$$

PROOF. We start by Case (i), and follow the lines of proof of Lemma 2.5 in [19]. First, we select a strictly increasing sequence $\{x_j : j \geqq 1\}$ with $x_j \geqq j$ for $j = 1, 2, \ldots$, and such that

$$\text{(2.31)} \qquad \lim_{j\to\infty}\Big\{-\frac{\log \mathbf{P}(X \geqq x_j)}{\psi(x_j)}\Big\} = 1,$$

which is rendered possible by Lemma 2.9. Then, we choose an $\varepsilon > 0$ and define a sequence of integers $\{n_j : j \geqq 1\}$ by

$$n_j = \max\Big\{n : \gamma\Big(\frac{\log k_n}{(1+\varepsilon)^3 n}\Big) \leqq x_j\Big\} \quad \text{for} \quad j \geqq 1.$$

Recalling from (1.9)–(1.10) that γ is a strictly increasing and continuous concave function on $[0,\infty)$ such that $\gamma(z) \to \infty$ as $z \to \infty$, we see that (2.20) implies the existence and finiteness of n_j for each $j \geqq 1$. Moreover, the fact that $x_j \to \infty$ implies that $n_j \to \infty$ as $j \to \infty$, so that, by eventually replacing $\{x_j : j \geqq 1\}$ by an appropriate subsequence, we may always assume that $n_j \geqq 1$ for $j \geqq 1$. Now, by (2.32), the continuity of γ implies, for each $j \geqq 1$, the existence of a θ_j such that

$$\text{(2.33)} \qquad \gamma\Big(\frac{\log k_{n_j}}{(1+\varepsilon)^3 n_j}\Big) \leqq x_j = \gamma\Big(\frac{(1+\theta_j)\log k_{n_j}}{(1+\varepsilon)^3 n_j}\Big) < \gamma\Big(\frac{\log k_{n_j+1}}{(1+\varepsilon)^3 (n_j+1)}\Big).$$

By (2.33), our assumptions and the fact that γ is strictly increasing jointly imply that

$$1 \leqq 1 + \theta_j \leqq \Big(\frac{n_j}{n_j+1}\Big)\frac{\log k_{n_j+1}}{\log k_{n_j}} \to 1 \quad \text{as} \quad j \to \infty,$$

so that $\theta_j \to 0$ as $j \to \infty$, and we have, for all large j, the inequalities

$$\text{(2.34)} \qquad 0 \leqq \theta_j \leqq \varepsilon.$$

Next, observe that the same arguments which have been used to establish (2.23)–(2.24) allow to prove that, for an arbitrary ζ and $n \geqq 1$,

$$\mathbf{P}(M_n(k_n) < n\zeta) = \mathbf{P}(S_{1,n} < n\zeta)^{k_n} \leqq \{1 - \mathbf{P}(X \geqq \zeta)^n\}^{k_n}.$$

This inequality, written with $n = n_j$ and $\zeta = \frac{1}{(1+\varepsilon)^3}\gamma\left(\frac{\log k_{n_j}}{n_j}\right)$, yields

$$\begin{aligned}(2.35)\qquad &\mathbf{P}\Big(M_{n_j}(k_{n_j}) < \frac{n_j}{(1+\varepsilon)^3}\gamma\Big(\frac{\log k_{n_j}}{n_j}\Big)\Big)\\ &\leqq \Big\{1-\mathbf{P}\Big(X \geqq \frac{1}{(1+\varepsilon)^3}\gamma\Big(\frac{\log k_{n_j}}{n_j}\Big)\Big)^{n_j}\Big\}^{k_{n_j}}.\end{aligned}$$

Since γ is concave on $[0,\infty)$ with $\gamma(0)=0$, we have $\lambda\gamma(z) \leqq \gamma(\lambda z)$ for all $z \geqq 0$ and $0 \leqq \lambda \leqq 1$. By choosing in this inequality $\lambda = \frac{1}{(1+\varepsilon)^3}$ and $z = \gamma\left(\frac{\log k_{n_j}}{n_j}\right)$, we readily infer from (2.33) that

$$\begin{aligned}(2.36)\qquad &\mathbf{P}\Big(X \geqq \frac{1}{(1+\varepsilon)^3}\gamma\Big(\frac{\log k_{n_j}}{n_j}\Big)\Big) \geqq \mathbf{P}\Big(X \geqq \gamma\Big(\frac{\log k_{n_j}}{(1+\varepsilon)^3 n_j}\Big)\Big)\\ &\geqq \mathbf{P}\Big(X \geqq \gamma\Big(\frac{(1+\theta_j)\log k_{n_j}}{(1+\varepsilon)^3 n_j}\Big)\Big) = \mathbf{P}(X \geqq x_j).\end{aligned}$$

Making use of (2.31), in combination with (2.33) and (2.34), we see that for all large j,

$$\begin{aligned}(2.37)\qquad &\mathbf{P}(X \geqq x_j) \geqq \exp(-(1+\varepsilon)\psi(x_j)) =\\ &= \exp\Big(-\frac{(1+\theta_j)\log k_{n_j}}{(1+\varepsilon)^2 n_j}\Big) \geqq \exp\Big(-\frac{\log k_{n_j}}{(1+\varepsilon)n_j}\Big).\end{aligned}$$

By combining (2.35), (2.36) and (2.37), we obtain that, for all j sufficiently large,

$$\begin{aligned}&\mathbf{P}\Big(M_{n_j}(k_{n_j}) < \frac{n_j}{(1+\varepsilon)^3}\gamma\Big(\frac{\log k_{n_j}}{n_j}\Big)\Big) \leqq \Big\{1-\exp\Big(-\frac{\log k_{n_j}}{1+\varepsilon}\Big)\Big\}^{k_{n_j}} \leqq\\ &\leqq \exp\Big(-\exp\Big(\frac{\varepsilon}{1+\varepsilon}\log k_{n_j}\Big)\Big),\end{aligned}$$

which, by (2.20) and using the fact that $n_j \geqq j$, is readily verified to be summable in j. An application of the Borel–Cantelli lemma, together with choosing $\varepsilon > 0$ arbitrarily small completes the proof of (2.30).

The proof of the lemma in Case (ii) is almost identical to the just-given proof given in Case (i), with the exception of the following details. Given an arbitrary $\varepsilon > 0$, we choose *first* a sequence $\{n_j : j \geqq 0\}$ of positive integers, with $n_j \uparrow \infty$, and such that, for all $j \geqq 1$, there exists an x_j satisfying $\mathbf{P}(X \geqq x_j) \geqq \exp(-(1+\varepsilon)\psi(x_j))$ and

$$\frac{1}{(1+\varepsilon)^3}\frac{\log k_{n_j}}{n_j} \leqq x_j \leqq \frac{1}{(1+\varepsilon)^2}\frac{\log k_{n_j}}{n_j}.$$

The existence of such a sequence is guaranteed by (H8). We then let θ_j be defined similarly as in (2.33), via the relations

$$\gamma\Big(\frac{\log k_{n_j}}{(1+\varepsilon)^3 n_j}\Big) \leqq x_j = \gamma\Big(\frac{(1+\theta_j)\log k_{n_j}}{(1+\varepsilon)^3 n_j}\Big) \leqq \gamma\Big(\frac{\log k_{n_j}}{(1+\varepsilon)^2 n_j}\Big).$$

Because of the above inequalities, (2.34) holds with this new definition of θ_j. Aside of this different construction of the sequences $\{x_j : j \geqq 1\}$ and $\{n_j : j \geqq 1\}$, the remainder of the proof follows along the same lines as that of Case (i). In particular, no change is needed in the relations (2.35)–(2.37). □

REMARK 2.4. (a) In view of (2.29), the assumption (H8) in Lemma 2.10 is satisfied under the weaker assumption that

$$\text{(H9)} \quad \lim_{n\to\infty}\Big(\sup_{\lambda\frac{\log k_n}{n}\leqq x\leqq\frac{\log k_n}{n}}\Big\{\frac{-\log(1-F(x))}{\psi(x)}\Big\}\Big) = 1, \text{ for some } \lambda\in(0,1).$$

The latter assumption is itself implied by (H4) of Theorem 1.4, which we recall below for convenience.

$$\text{(H4)} \qquad (-\log(1-F(x)))/\psi(x) \to 1 \quad \text{as} \quad x\to\infty.$$

As shown by Mason [19], the assumption (H4) implies (H5), that is,

$$\text{(H5)} \qquad \lim_{x\to\infty} x^{-1}\gamma(-\log(1-F(x))) = 1,$$

whereas the converse is not true.

(b) The condition $(\log k_{n+1})/\log k_n \to 1$ is satisfied by all sequences of the form $k_n = \exp(n^\alpha)$ with $\alpha > 0$, but not by $k_n = \exp(e^{n\rho})$ with $\rho > 0$.

(c) The general meaning of either of the assumptions (H8), (H9) or $(\log k_{n+1})/\log k_n \to 1$ is that it is possible to find a sequence $x_j \to \infty$ along which $(-\log(1-F(x))/\psi(x) \to 1$, which is *sufficiently close* to $n_j^{-1}\log k_{n_j}$ for an appropriate choice of the n_j's. We will not investigate the case when this property does not hold, since it corresponds either to *unseldom distributions* (which do not satisfy (H4)), or *very sparse* sequences (such as $k_n = \exp(n^\rho)$).

LEMMA 2.11. *Under (H1)-(H2)-(H5), assume that* (2.20) *holds, and that* $\omega = \infty$. *Then*

$$\text{(2.38)} \qquad \liminf_{n\to\infty}\Big\{\frac{M_n(k_n)}{n\gamma(n^{-1}\log k_n)}\Big\} \geqq 1 \quad a.s..$$

PROOF. Recall from (H5) that $z^{-1}(-\log(1-F(z))) \to 1$ as $z\to\infty$. Fix an arbitrary $\varepsilon > 0$. It follows from (2.20) that, for all large n,

$$\gamma\Big(-\log\Big(1-F\Big(\frac{1}{(1+\varepsilon)^2}\gamma\Big(\frac{\log k_n}{n}\Big)\Big)\Big)\Big) < \frac{1}{1+\varepsilon}\gamma\Big(\frac{\log k_n}{n}\Big).$$

Making use of the concavity of γ, which implies that $\lambda\gamma(x) \leqq \gamma(\lambda x)$ for $0 \leqq \lambda \leqq 1$ and $x \geqq 0$, we infer from this last inequality that

$$\begin{aligned}\mathbf{P}\Big(X > \frac{1}{(1+\varepsilon)^2}\gamma\Big(\frac{\log k_n}{n}\Big)\Big) &\geqq \exp\Big(-\psi\Big(\frac{1}{1+\varepsilon}\gamma\Big(\frac{\log k_n}{n}\Big)\Big)\Big) \\ &\geqq \exp\Big(-\psi\Big(\gamma\Big(\frac{\log k_n}{(1+\varepsilon)n}\Big)\Big)\Big) = \exp\Big(-\frac{\log k_n}{(1+\varepsilon)n}\Big).\end{aligned} \tag{2.39}$$

Given (2.39), the remainder of the proof is very similar to (2.35)–(2.37) in the proof of Lemma 2.10. Therefore, we omit the details. □

PROOF OF THEOREM 1.4. In view of Remark 2.4(a), it is readily obtained by combining Lemmas 2.6, 2.8, 2.10 and 2.11. We note that the conclusion of the theorem holds under the formal replacement of (H4) by either (H8) or (H9). □

When (H5) does not hold, Theorem 1.4 does not give any information on the liminf behaviour of $M_n(k_n)$. Our next theorem, proved under very general assumptions, fills in part this gap. Introduce the following notation. Define the *quantile function* pertaining to the d.f. $F(x) = \mathbf{P}(X \leqq x)$ by $Q(s) = \inf\{x : F(x) \geqq s\}$ for $0 < s < 1$.

THEOREM 2.2. *Let $\{k_n : n \geqq 1\}$ be an arbitrary sequence such that*

$$\delta := \liminf_{n\to\infty}\Big\{\frac{k_n}{\log n}\Big\} > 1. \tag{2.40}$$

Then, for any $0 < \varepsilon < \delta - 1$, we have

$$\liminf_{n\to\infty}\Big\{\frac{M_n(k_n)}{nQ(1-\exp(-n^{-1}(\log k_n - \log_2 n^{1+\varepsilon})))}\Big\} \geqq 1 \quad a.s.. \tag{2.41}$$

PROOF. Without loss of generality, we may and do assume that, for $i \geqq 1$ and $j \geqq 1$, $X_{i,j} = Q(U_{i,j})$, where $\{U_{i,j} : i \geqq 1, j \geqq 1\}$ is an array of i.i.d. uniform $(0,1)$ random variables. Clearly, for each $n \geqq 1$, we have $M_n(k_n) \geqq nQ(V_n)$, where we set

$$V_n = \max_{1\leqq i\leqq k_n}\Big(\min_{1\leqq j\leqq n} U_{i,j}\Big).$$

Fix any $\varepsilon \in (0, d)$. Since, for $0 < u < 1$, $\mathbf{P}(V_n \leqq 1-u) = (1-u^n)^{k_n}$, by setting

$$u = u_n(\varepsilon) := \exp(-n^{-1}(\log k_n - \log_2 n^{1+\varepsilon})),$$

we see that for all large n,

$$\mathbf{P}(V_n \leqq 1 - u_n(\varepsilon)) = \Big(1 - \frac{(1+\varepsilon)\log n}{k_n}\Big)^{k_n} \leqq \frac{1}{n^{1+\varepsilon}}. \tag{2.42}$$

Here, we have made use of (2.40) and $1+\varepsilon<d$ which jointly imply that ultimately $0<\frac{(1+\varepsilon)\log n}{k_n}<1$. Since the right side of (2.42) is summable, (2.41) follows by an application of the Borel–Cantelli lemma. $\square$

REMARK 2.5. Assume that (H1)–(H2)–(H3) hold, and that $\omega=\infty$. By Lemma 2.6 of [19], (H5) is then equivalent to $\gamma(\log n)/Q(1-1/n)\to 1$ as $n\to\infty$. The concavity and monotoneness of γ show in turn that the latter condition is equivalent to

$$\text{(H10)} \qquad \gamma(u)/Q(1-\exp(-u))\to 1 \quad \text{as} \quad u\to\infty.$$

Set $u_n(\varepsilon)=\exp(-n^{-1}(\log k_n-\log_2 n^{1+\varepsilon}))$ and $v_n=\exp(-n^{-1}\log k_n)$. It is easily checked that, for any $\varepsilon>0$, $u_n(\varepsilon)/v_n\to 1$. Therefore, whenever $\liminf_{n\to\infty}\{n^{-1}\log k_n\}>0$, we have $\gamma(u_n(\varepsilon))/\gamma(v_n)\to 1$. This, in turn, shows that, under (H1)–(H2)–(H3)–(H5) and $\liminf_{n\to\infty}\{n^{-1}\log k_n\}>0$, the conclusion (2.41) of Theorem 2.2 is equivalent to

$$\liminf_{n\to\infty}\left\{\frac{M_n(k_n)}{nQ(1-\exp(-n^{-1}\log k_n))}\right\}\geq 1 \quad \text{a.s..} \tag{2.43}$$

This is in agreement with Theorem 1.4, since, under (H10), the conclusion (1.15) of this theorem is equivalent to

$$\lim_{n\to\infty}\left\{\frac{M_n(k_n)}{nQ(1-\exp(-n^{-1}\log k_n))}\right\}=1 \quad \text{a.s..} \tag{2.44}$$

REMARK 2.6. If $n^{-1}\log k_n\to\infty$, the equivalence between (2.41) and (2.43) may hold under very general assumptions without assuming either the finiteness of moments of X, or that (H1)–(H3) hold. For example, it is readily checked that (2.41) and (2.43) are equivalent under the condition

$$\text{(H11)} \qquad \lim_{\lambda\downarrow 1}\left(\limsup_{s\downarrow 0}\left\{\frac{Q(1-s)}{Q(1-\lambda s)}\right\}\right)=1,$$

which is satisfied, in particular, when the upper tail of X is in the domain of attraction of an *extreme-value distribution* (see e.g. Chapter 2 of de Haan [13]).

2.3. Moderate to small sequences. In this last subsection, we consider results which hold under the moment-type assumptions (H7)–(H6) imposed upon X, and give successively the proofs of Theorems 1.6 and 1.5. Our arguments will combine invariance principles, together with results of Deheuvels and Teicher [7], the latter being applied to partial sums of i.i.d. $N(0,1)$ random variables.

First, we recall from the invariance principles of Komlós, Major and Tusnády [14], Major [18] and Einmahl [9], [10], the following basic fact. Let H be a positive nondecreasing function on $(0,\infty)$, such that

$$\text{(2.45)}\qquad \begin{aligned} &\mathbf{E}(H(X))<\infty; &&\frac{H(x)}{x^2\log_2 x}\uparrow \text{ eventually;}\\ &\frac{\log H(x)}{x}\downarrow \text{ eventually;} &&\liminf_{x\uparrow\infty}\frac{H(\varepsilon x)}{H(x)}>0 \text{ for all } \varepsilon>0. \end{aligned}$$

Set further $H^{-1}(x)=\inf\{s\geqq 0\colon H(s)\geqq x\}$ for $x\geqq 0$. Then, on a suitable probability space, it is possible to construct the random variables $\{X_{i,j}\colon i\geqq 1,\ j\geqq 1\}$ together with a sequence of i.i.d. standard Wiener processes $\{W_i(t)\colon i\geqq 1,\ t\geqq 0\}$, in such a way that

$$\text{(2.46)}\qquad \max_{1\leqq i\leqq k_n}|S_{i,n}-W_i(n)|=o(H^{-1}(nk_n))\quad \text{a.s. as } n\to\infty.$$

Proof of Theorem 1.6. Given a sequence of i.i.d. standard Wiener processes $\{W_i(t)\colon i\geqq 1,\ t\geqq 0\}$, we first set $Y_{i,j}=W_i(j)-W_i(j-1)$ for $i\geqq 1$ and $j\geqq 1$, and observe that $\{Y_{i,j}\colon i\geqq 1,\ j\geqq 1\}$ so defined constitute an array of i.i.d. $N(0,1)$ random variables. Set $Y=Y_{1,1}$ and

$$\text{(2.47)}\qquad M_n^*(k_n)=\max_{1\leqq i\leqq k_n}W_i(n).$$

By applying Theorem 5.4 of [7] to $\{Y_{i,j}\colon i\geqq 1,\ j\geqq 1\}$ we see that, whenever $\{k_n\colon n\geqq 1\}$ is such that $k_n\uparrow$, $n^{-1}\log k_n\downarrow 0$, $(\log k_n)/\log_2 n\to\infty$, and

$$\text{(2.48)}\qquad \sum_{n=1}^{\infty}k_n\mathbf{P}\Big(|Y|>\varepsilon\Big(\frac{n}{\log k_n}\Big)^{1/2}\Big)<\infty,$$

we have

$$\text{(2.49)}\qquad \lim_{n\to\infty}(2n\log k_n)^{-1/2}M_n^*(k_n)=1\quad \text{a.s..}$$

In view of the well-known fact that

$$\text{(2.50)}\qquad \mathbf{P}(|Y|>x)=2(1-\Phi(x))=\frac{1+o(1)}{x\sqrt{2\pi}}e^{-x^2/2}\quad \text{as } x\to\infty,$$

the assumptions upon $\{k_n\colon n\geqq 1\}$ made in Theorem 1.5 imply that

$$k_n\mathbf{P}\Big(|Y|>\varepsilon\Big(\frac{n}{\log k_n}\Big)^{1/2}\Big)=O\Big(k_n\Big(\frac{\log k_n}{n}\Big)^{1/2}\exp\Big(-\frac{n\varepsilon^2}{2\log k_n}\Big)\Big),$$

which, by (1.17), is summable, implying therefore that (2.48) and (2.49) hold.

Next, we check readily from (H7) that the function

$$H(x) = x^{2\alpha+2}/(\log_+ x)^{\alpha+1}$$

satisfies (2.45). Since then

$$H^{-1}(y) = O(y^{\frac{1}{2\alpha+2}}(\log y)^{1/2}) \quad \text{as} \quad y \to \infty,$$

it follows from (1.17) that $H^{-1}(nk_n) = O\{(n \log n)^{1/2}\}$, whence we may infer (1.4) from (2.46) and (2.49). □

REMARK 2.7. Theorem 1.6 improves upon Corollary 5.1 of Deheuvels and Teicher (see e.g. Theorem 1.2). The following theorem shows that the moment assumption (H5) in Theorem 1.6 is sharp.

THEOREM 2.3. *Let $k_n = n^\alpha$ for some $\alpha > 0$. Then, a necessary condition for*

$$\limsup_{n\to\infty} (2n \log k_n)^{-1/2} \max_{1 \leqq i \leqq k_n} |S_{i,n}| < \infty \quad a.s.,$$

is that (H5) holds, i.e.

$$\mathbf{E}\{|X|^{2\alpha+2}/(\log_+ |X|)^{\alpha+1}\} < \infty.$$

PROOF. Choose any $\lambda \in (0, 1)$, and set $m_n = \lambda(2n \log k_n)^{1/2}$, with $k_n = n^\alpha$ in Lemma 2.5. Since then (2.8) is satisfied, we have (2.10), i.e.

$$\sum_{j=1}^{\infty} \mathbf{P}(dm_j \leqq |X| < dm_{j+1}) \sum_{n=1}^{j} k_n < \infty.$$

Given this fact, the observation that, for some $c_0, c_1 > 0$ and all large j,

$$\sum_{n=1}^{j} k_n \geqq c_1 j^{\alpha+1} \geqq c_0 m_{j+1}^{2\alpha+2} (\log m_{j+1})^{-\alpha+1},$$

readily implies (H5). □

PROOF OF THEOREM 1.5. We follow the same steps as in the proof of Theorem 1.6, with the following modifications. First, we choose $H(x) = \exp(\lambda x^{2\beta/(1+\beta)})$, which, in view of (H4), obviously satisfies (2.45). Given this choice, we see that

$$H^{-1}(y) = \left\{\frac{1}{\lambda} \log y\right\}^{\frac{1+\beta}{2\beta}} \quad \text{for} \quad y > 0. \tag{2.51}$$

It follows from (2.51) that, under (1.16), $H^{-1}(nk_n)=O(n^{\frac{1}{2}(1+\beta)})$. Moreover, (1.16) implies that, as $n\to\infty$, $(2n\log k_n)^{1/2}=(1+o(1))\rho^{1/2}n^{\frac{1}{2}(1+\beta)}$. Thus, letting $M_n^*(k_n)$ be as in (2.47), we may infer from (1.46) that

$$(2n\log k_n)^{-1/2}|M_n^*(k_n)-M_n(k_n)|\to 0 \quad \text{a.s. as } n\to\infty.$$

It suffices therefore to establish (1.4) with the formal replacement of $M_n(k_n)$ by $M_n^*(k_n)$. Towards this aim, we fix $\varepsilon>0$ and use the Bonferroni inequality together with (2.50) to obtain the simple bound

$$\begin{aligned}\mathbf{P}(M_n^*(k_n)>(2(1+\varepsilon)n\log k_n)^{1/2})&\leqq k_n\mathbf{P}(n^{-1/2}W_1(n)>(2(1+\varepsilon)\log k_n)^{1/2})=\\ &=O\Big(\frac{k_n}{\sqrt{\log k_n}}\exp(-(1+\varepsilon)\log k_n)\Big)=O(k_n^{-\varepsilon}),\end{aligned}$$

which, in view of (1.16), is summable in n. Therefore, an application of the Borel–Cantelli lemma, together with the fact that $\varepsilon>0$ may be chosen arbitrarily small, shows that

$$\limsup_{n\to\infty}(2n\log k_n)^{-1/2}M_n(k_n)\leqq 1 \quad \text{a.s.}. \tag{2.52}$$

To obtain the reverse inequality, we observe that, for any $0<\varepsilon<1$, we have, for some constants $c_2,c_3>0$ and all large n

$$\begin{aligned}\mathbf{P}(M_n^*(k_n)&\leqq(2(1-\varepsilon)n\log k_n)^{1/2})=\\ &=\{1-\mathbf{P}(n^{-1/2}W_1(n)>(2(1-\varepsilon)\log k_n)^{1/2})\}^{\lfloor k_n\rfloor}\\ &\leqq\exp\Big(-c_2\frac{k_n}{\sqrt{\log k_n}}e^{-(1-\varepsilon)\log k_n}\Big)\\ &=O\Big(\exp\Big(-c_3\frac{k_n^{\varepsilon}}{\sqrt{\log k_n}}\Big)\Big),\end{aligned}$$

which is summable in n. Since $\varepsilon>0$ is arbitrary, we conclude from the Borel–Cantelli lemma that

$$\liminf_{n\to\infty}(2n\log k_n)^{-1/2}M_n(k_n)\geqq 1 \quad \text{a.s.}. \tag{2.53}$$

The proof is completed by combining (2.52) with (2.53). □

The following theorem shows that the moment assumption (H4) in Theorem 1.5 is sharp.

THEOREM 2.4. *Let $k_n=\exp(n^{\beta})$ for some $\beta\in(0,1)$. Then, a necessary condition for*

$$\limsup_{n\to\infty}(2n\log k_n)^{-1/2}\max_{1\leqq i\leqq k_n}|S_{i,n}|<\infty \quad a.s.,$$

is that (H4) holds, i.e.

$$\mathbf{E}\{\exp(\lambda|X|^{\frac{2\beta}{\beta+1}})\} < \infty \quad \textit{for some} \quad \lambda > 0.$$

Proof. Set $m_n = (2n\log k_n)^{1/2}$, with $k_n = \exp(n^\beta)$ in Lemma 2.5. Since then (2.8) is satisfied, we have (2.10), i.e.

$$\sum_{j=1}^{\infty} \mathbf{P}(dm_j \leqq |X| < dm_{j+1}) \sum_{n=1}^{j} k_n < \infty. \tag{2.54}$$

Given this fact, we observe that, for some $c_4, c_5 > 0$ and all large j,

$$\sum_{n=1}^{j} k_n \geqq c_4 j \exp(c_5 j^\beta). \tag{2.55}$$

Since $m_j = (1+o(1))2^{1/2} j^{\frac{1}{2}(1+\beta)}$, we have, for some suitable $c_6, c_7 > 0$ and all large j, $c_4 j \geqq c_6 m_{j+1}^{2/(1+\beta)}$, $c_5 j^\beta \geqq 2c_7 m_{j+1}^{2\beta/(1+\beta)}$ and

$$c_4 j \exp(c_5 j^\beta) \geqq \exp(c_7 m_{j+1}^{2\beta/(1+\beta)}).$$

By combining this last inequality with (2.54) and (2.55), we readily obtain that (H4) holds with $\lambda = c_7$. □

REFERENCES

[1] Acosta, A. de and Kuelbs, J., Limit theorems for moving averages of independent random vectors, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **64** (1983), 67–123. *MR* **84i**:60011

[2] Bártfai, P., Die Bestimmung der zu einem wiederkehrenden Prozess gehörenden Verteilungsfunktion aus den mit Fehlern behafteten Daten einer einzigen Realisation, *Studia. Sci. Math. Hungar.* **1** (1966), 161–168. *MR* **35** #6218

[3] Chernoff, H., A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations, *Ann. Math. Statist.* **23** (1952), 493–507. *MR* **15**–241

[4] Cramér, H., Su un teorema relativo alla legge uniforme dei grandi numeri, *Giorn. Ist. Ital. Attuari* **5** (1934), 1–13. *Zbl* **8**, 368

[5] Deheuvels, P., Functional Erdős–Rényi laws, *Studia Sci. Math. Hungar.* **26** (1991), 261–295. *MR* **93i**:60056

[6] Deheuvels, P. and Vervaat, W., Large deviations for rocket plumes, 1994 (unpublished manuscript).

[7] Deheuvels, P. and Teicher, H., A new law of the iterated logarithm for arrays, *Studia Sci. Math. Hungar.* **31** (1996), 145–185.

[8] Deheuvels, P. and Révész, P., On the coverage of Strassen-type sets by sequences of Wiener processes, *J. Theoret. Probab.* **6** (1993), 427–449. *MR* **94k**:60063

[9] Einmahl, U., A useful estimate in the multidimensional invariance principle, *Probab. Theory Related Fields* **76** (1987), 81–101. *MR* **88k**:60063

[10] Einmahl, U., Strong invariance principles for partial sums of independent random vectors, *Ann. Probab.* **15** (1987), 1419–1440. *MR* **88h**:60071

[11] Erdős, P. and Rényi, A., On a new law of large numbers, *J. Analyse Math.* **23** (1970), 103–111. *MR* **42** #6907

[12] Esséen, C.-G., Fourier analysis of distribution functions. A mathematical study of the Laplace–Gaussian law, *Acta. Math.* **77** (1945), 1–125. *MR* **7**–312

[13] Haan, L. de, On regular variation and its application to the weak convergence of sample extremes, *Math. Centre Tracts* **32**, Mathematisch Centrum, Amsterdam, 1975.

[14] Komlós, J., Major, P. and Tusnády, G., An approximation of partial sums of independent RV's and the sample DF. II, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **34** (1976), 33–58. *MR* **53** #6697

[15] LePage, R. and Schreiber, B. M., A square root law for diffusing particles, *Probability in Banach spaces, V* (Medford, Mass., 1984), Lecture Notes in Mathematics, **1153**, Springer-Verlag, Berlin–New York, 1985, 323–328. *MR* **87j**:60054

[16] LePage, R. and Schreiber, B. M., An iterated logarithm law for families of Brownian paths, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **70** (1985), 341–344. *MR* **86m**:60198

[17] LePage, R. and Schreiber, B. M., Addendum to "An iterated logarithm law for families of Brownian paths", *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **76** (1987), 629. *MR* **88m**:60206

[18] Major, P., The approximation of partial sums of independent RV's, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **35** (1976), 213–220. *MR* **54** #3823

[19] Mason, D. M., An extended version of the Erdős–Rényi strong law of large numbers, *Ann. Probab.* **17** (1989), 257–265. *MR* **91a**:60082

[20] Mogulskii, A. A., Large deviations for processes with independent increments, *Ann. Probab.* **21** (1993), 202–215. *MR* **94g**:60053

[21] Schilder, M., Some asymptotic formulae for Wiener integrals, *Trans. Amer. Math. Soc.* **125** (1966), 63–85. *MR* **34** #1770

[22] Steinebach, J., On a necessary condition for the Erdős–Rényi law of large numbers, *Proc. Amer. Math. Soc.* **68** (1978), 97–100. *MR* **57** # 1622

(Received October 24, 1994)

UNIVERSITÉ PIERRE ET MARIE CURIE
LABORATOIRE DE STATISTIQUE
THÉORIQUE ET APPLIQUÉE
TOUR 45–55-3^e–BOÎTE NO. 158
4, PLACE JUSSIEU
F–75252 PARIS Cedex 05
FRANCE

pd@ ccr.jussieu.fr

FACHBEREICH MATHEMATIK
PHILIPPS UNIVERSITÄT
D–35032 MARBURG
FEDERAL REPUBLIC OF GERMANY

jost@ mathematik.uni-marburg.de

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 145–185

A NEW LAW OF THE ITERATED LOGARITHM FOR ARRAYS

P. DEHEUVELS and H. TEICHER

Dedicated to Pál Révész for his sixtieth birthday

Abstract

Let $\mathcal{X}=\{X_{ij},\ i\geqq 1, j\geqq 1\}$ denote an array of centered random variables with finite second moment, and set $S_{in}=\sum_{j=1}^{n} X_{ij}$ for $i\geqq 1$ and $n\geqq 1$. Let $M_{k_n}(n)=\max\limits_{1\leqq i\leqq k_n} S_{in}$, where $k_n\geqq 1$ is a nondecreasing sequence of real numbers. Under suitable conditions upon the distribution of $\mathcal{X}$ and on the growth of k_n, we characterize the almost sure limit set of $b_n^{-1}M_{k_n}(n)$, where b_n is an appropriate sequence of norming constants, depending upon n and k_n. For $k_n=1$, our results include as a particular case the law of the iterated logarithm for sums of non-identically distributed random variables.

1. Introduction

Let $\{\Omega,\mathcal{A},\mathbf{P}\}$ be a probability space on which is defined an array $\mathcal{X}=$ $=\{X_{ij},\ i\geqq 1, j\geqq 1\}$ of centered random variables with finite second moment, and let $S_{in}=\sum_{j=1}^{n} X_{ij}$ for $i\geqq 1$ and $n\geqq 1$. The question to be explored is the limiting behaviour of $M_{k_n}=M_{k_n}(n)=\max\limits_{1\leqq i\leqq k_n} S_{in}$ as n tends to infinity, where $\{k_n, n\geqq 1\}$ is a non-decreasing sequence of real numbers with $k_n\geqq 1$ for $n\geqq 1$. As it will be shown, the nature of this behaviour depends quintessentially on the magnitude of k_n. The underlying assumptions on $\mathcal{X}$ will be either (as in Section 5) that the component random variables are *totally independent and identically distributed* [t.i.i.d.] or more generally (as in Sections 3 and 4) that the component random variables are *totally independent* [t.i.], and *columnwise identically distributed* [c.i.d.], meaning that the distribution of the *column sequence* $\{X_{in}, n\geqq 1\}$ is independent of $i\geqq 1$.

To motivate the study of $M_{k_n}(n)$, we start by considering some examples of the results we have in mind. Let $\mathcal{X}$ be t.i.i.d., and denote by X a random variable following the common distribution of X_{ij} for $i\geqq 1$ and $j\geqq 1$. Assume that

$$\mathbf{E}X=0 \quad\text{and}\quad \mathbf{E}X^2=1. \tag{1.1}$$

1991 *Mathematics Subject Classification.* Primary 60F15; Secondary 60G15.

Key words and phrases. Law of the iterated logarithm, double arrays, almost sure convergence, strong laws.

0081-6906/96/$ 5.00 ©1996 *Akadémiai Kiadó, Budapest*

Under (1.1) and for $k_n = 1$ the strong limiting behaviour of $M_1(n) = S_{1n}$ is described by the *law of the iterated logarithm* [LIL] (Hartman and Wintner [17]) which asserts that

$$\limsup_{n\to\infty} \pm(2nLLn)^{-1/2} S_{1n} = 1 \quad \text{a.s.,} \tag{1.2}$$

where we set $Lu = \log(u \vee e)$, $LLu = L(Lu)$, $L_0 u = u$ and $L_q u = L(L_{q-1}u)$ for $q \geqq 1$. Moreover, the *limit set* [i.e. the set of limit points in $\mathbf{R}$ endowed with the usual topology] of the sequence $\{(2nLLn)^{-1/2} M_1(n), n \geqq 1\}$ is almost surely equal to the interval $[-1, 1]$ [26]. For a fixed $k_n = k \geqq 1$, the following extension of (1.2), due to Finkelstein [14] holds.

Lemma 1.1. *Let $\mathcal{X}$ be t.i.i.d. and let $k \geqq 1$ be a fixed integer. Then, under* (1.1), *the limit set in $\mathbf{R}^k$ of the sequence $\{((2nLLn)^{-1/2} S_{1n}, \dots, (2nLLn)^{-1/2} S_{kn}),\ n \geqq 1\}$ is almost surely equal to the unit ball of the Euclidean space $\mathbf{R}^k$ denoted by $\mathcal{B}_k = \{(z_1, \dots, z_k) \in \mathbf{R}^k : \sum_{j=1}^k z_j^2 \leqq 1\}$.*

The following theorem is a consequence of Lemma 1.1 (see Lemma 2.2 in the sequel).

Theorem 1.1. *Let $\mathcal{X}$ be t.i.i.d. and let $k \geqq 1$ be a fixed integer. Then, under* (1.1), *the limit set of the sequence $\{(2nLLn)^{-1/2} M_k(n), n \geqq 1\}$ is almost surely equal to the interval $[-k^{-1/2}, 1]$.*

In Section 2, we will prove an extended version of Theorem 1.1, together, with a generalization of Finkelstein's Lemma 1.1 and the consequences thereof for the non-identically distributed case. Other useful preliminaries will also be treated in this section.

In view of Theorem 1.1, it is natural to seek a characterization of the *almost sure limit set* [i.e. the set of almost sure limit points] of the sequence $(2nLLn)^{-1/2} M_{k_n}(n)$ when $k_n \to \infty$. An outline of our main results concerning this question is as follows. Assuming $\mathbf{E}X = 0$, $\mathbf{E}X^2 = 1$ and $\mathbf{E}X^2 g(|X|) < \infty$ for some appropriate slowly varying function g, we will prove that the almost sure limit set of $(2nLLn)^{-1/2} M_{k_n}(n)$ is equal to $[0, 1]$ when k_n is *"small"*, but shifts to $[\beta_1^{1/2}, (1 + \beta_1)^{1/2}]$ when k_n is of *"moderate"* size, such as, for $k_n = (Ln)^{\beta_1} (LLn)^{\beta_2}$ with $\beta_1 > 0$. On the other hand, when k_n is "large", the normalization changes and $(2nLk_n)^{-1/2} M_{k_n}(n) \to 1$ almost surely. The precise definition of *small*, *moderate* and *large* is given in Section 3, where we investigate the *upper limit* of $M_{k_n}(n)$. Section 4 deals with the *lower limit* and verifies that the interval between the lower and upper limits comprises the almost sure limit set. Section 5 simplifies and sharpens these results in the t.i.i.d. case.

In the context of a sequence of independent Wiener process results akin to those of the t.i.i.d. case for "moderate" and "small" k_n have been obtained by Le Page and Schreiber [21] and Deheuvels and Révész [8]. We mention in Section 6 the connection of these results with our theorems, via invariance principles. Especially, we will show how the strong approximation theorem of Komlós, Major and Tusnády [18], [19] and the *Skorokhod embedding scheme* for partial sums (Skorokhod [24]) may be used to reduce our LIL-type problem to that of establishing strong laws of large numbers [LLN] for arrays of auxiliary random variables.

The behaviour of S_{nn} has been investigated as early as 1934. In the t.i.i.d. case under (1.1) and $\mathbf{E}X^4 < \infty$, Cramér [3] and Esseen [11] have given a characterization of the upper and lower class of $n^{-1/2}S_{nn}$ in terms of the convergence or the divergence of the series $\sum \lambda_n^{-1} \exp\left(-\frac{1}{2}\lambda_n^2\right)$. By applying their results to $\lambda_n = (2(1 \pm \varepsilon)Ln)^{1/2}$ with $\varepsilon > 0$ chosen arbitrarily small, we obtain that

$$\limsup_{n\to\infty}(2nLn)^{-1/2}S_{nn} = 1, \quad \text{a.s.}. \tag{1.3}$$

Since $M_{k_n}(n) \geqq S_{nn}$ when $k_n = n$, (1.3) confirms the fact for "large" k_n the normalizing constant of $M_{k_n}(n)$ no longer contains an iterated logarithm (see Remark 5.3 in the sequel).

Other works dealing with or related to limit laws for arrays are those of Baxter [1], Cramér [3], Csörgő and Hall [5], Dabrowski, Dehling and Philipp [6], Deheuvels and Steinebach [7], Eicker [9], Feller [12], Esseen [11], Gut [15], Lai and Wei [20], Smythe [25], Teicher [29], [30] and Tomkins [31].

2. Preliminaries and notations

A propos of the array $\mathcal{X} = \{X_{ij},\ i \geqq 1, j \geqq 1\}$, we will consider several alternative sets of assumptions as reflected in the following terminology. $\mathcal{X}$ will be said to be *columnwise independent* [c.i.] (respectively, *columnwise identically distributed* [c.i.d.] or *columnwise independent and identically distributed* [c.i.i.d.]), according as $\{(X_{i1}, X_{i2}, \dots), i \geqq 1\}$ are independent (respectively, identically distributed or independent and identically distributed) random vectors. Moreover, $\mathcal{X}$ will be called *totally independent* [t.i.] (respectively *totally independent and identically distributed* [t.i.i.d.] if $\{X_{11}, X_{12}, X_{21}, X_{31}, X_{22}, X_{13}, \dots\}$ is a sequence of independent (resp. i.i.d.) random variables. Finally, $\mathcal{X}$ is *identically distributed* [i.d.] if all random variables in the array have a common distribution. In such a case, X will denote a generic random variable with this distribution.

In what follows, $I[A]$ denotes the indicator function of the set (or event) A, Z' signifies the transpose of the vector Z, and $[u]$ denotes the integer part of u. The first result, stated in Lemma 2.1 below, is a generalization of Lemma 2 of Finkelstein [14], stated in Lemma 1.1.

LEMMA 2.1. *Let* $Z_j' = (X_{1k}, \dots, X_{kj})$, $j \geqq 1$ *be independent random vectors whose components are i.i.d. with* $\mathbf{E}X_{1j} = 0$ *and* $\mathbf{E}X_{1j}^2 = \sigma_j^2 < \infty$ *for* $j \geqq 1$.

Assume further that

$$(2.1)\qquad s_n^2=\sum_{j=1}^{n}\sigma_j^2\to\infty,\qquad \sigma_n^2=o(s_n^2/LLs_n^2)\quad as\quad n\to\infty,$$

and, for some $\alpha\in(1,2]$ *and all* $\varepsilon>0$,

$$(2.2)\qquad \sum_{j=1}^{n}\mathbf{E}X_{1j}^2I\left[|X_{1j}|>\varepsilon s_j(LLs_j^2)^{-1/2}\right]=o(s_n^2),$$

$$(2.3)\qquad \sum_{n=1}^{\infty}\frac{1}{s_n^{\alpha}(LLs_n^2)^{\alpha/2}}\mathbf{E}|X_{1n}|^{\alpha}I\left[|X_{1n}|>\varepsilon s_n(LLs_n^2)^{-1/2}\right]<\infty.$$

Let $\Sigma_n=(2s_n^2LLs_n^2)^{-1/2}\sum_{j=1}^{n}Z_j$. *Then, the sequence* $\{\Sigma_n,n\geqq 1\}$ *is almost surely relatively compact in* $\mathbf{R}^k$ *with limit set equal to the unit ball* $\mathcal{B}_k=$ $=\left\{(z_1,\dots,z_k)\in\mathbf{R}^k:\sum_{j=1}^{k}z_j^2\leqq 1\right\}$ *of the Euclidean space* $\mathbf{R}^k$.

PROOF. Let $||\mathbf{x}||$ denote the Euclidean norm of $\mathbf{x}=(x_1,\dots,x_k)'\in\mathbf{R}^k$. The first step is to show, for any non-zero vector $\mathbf{t}=(t_1,\dots,t_k)'$, that the random variables $Y_j=\sum_{i=1}^{k}t_iX_{ij}$, $j\geqq 1$, obey

$$(2.4)\qquad \limsup_{n\to\infty}(2s_n^2LLs_n^2)^{-1/2}\sum_{j=1}^{n}Y_j=||\mathbf{t}||\quad \text{a.s.}.$$

According to the Corollary of Teicher [27] (see e.g. Corollary 10.2.4, p. 359 of Chow and Teicher [2]) (2.2) and (2.3) ensure (2.4) when $t_1=1$ and $t_i=0$ for $2\leqq i\leqq k$. Since $\{Y_j,j\geqq 1\}$ are independent random variables with $\mathbf{E}Y_j=0$, $\mathbf{E}Y_j^2=\sigma_j^2||\mathbf{t}||^2$ and $s_n^2||\mathbf{t}||^2=\sum_{j=1}^{n}\mathbf{E}Y_j^2=:\hat{s}_n^2$, the aforementioned corollary will yield (2.4) once it is demonstrated that (2.2) and (2.3) hold when X_{1n} and s_n^2 are replaced respectively by Y_n and $\hat{s}_n^2$. Towards this end, note that

$$\frac{1}{\hat{s}_n^2}\sum_{j=1}^{n}\mathbf{E}Y_j^2I\left[Y_j^2>\varepsilon^2\hat{s}_j^2(LL\hat{s}_j^2)^{-1}\right]$$

$$\leqq\frac{1}{\hat{s}_n^2}\sum_{j=1}^{n}\mathbf{E}\left(||\mathbf{t}||^2\sum_{i=1}^{k}X_{ij}^2I\left[\sum_{i=1}^{k}X_{ij}^2>\varepsilon^2s_j^2(LL\hat{s}_j^2)^{-1}\right]\right)$$

$$\leqq\frac{k}{s_n^2}\sum_{j=1}^{n}\mathbf{E}X_{1j}^2I\left[\max_{1\leq i\leq k}X_{ij}^2>\varepsilon^2s_j^2(kLL\hat{s}_j^2)^{-1}\right]$$

$$\leqq \sum_{i=1}^{k} \frac{k}{s_n^2} \sum_{j=1}^{n} \mathbf{E} X_{1j}^2 I\left[X_{ij}^2 > \varepsilon^2 s_j^2 (kLL\hat{s}_j^2)^{-1}\right]$$

$$= o(1) + \frac{k(k-1)}{s_n^2} \sum_{j=1}^{n} \sigma_j^2 \mathbf{P}\left(X_{2j}^2 > \varepsilon^2 s_j^2 (kLL\hat{s}_j^2)^{-1}\right) \to 0 \text{ as } n \to \infty,$$

where we have used (2.2) in combination with the fact, by Chebyshev's inequality, that

$$\mathbf{P}\left(X_{1j}^2 > \varepsilon^2 s_j^2 (kLL\hat{s}_j^2)^{-1}\right) \leqq \frac{k\sigma_j^2 LL\hat{s}_j^2}{\varepsilon^2 s_j^2} = o(1) \text{ as } j \to \infty.$$

In similar fashion, noting that for $\alpha > 1$

$$\left|\sum_{i=1}^{k} t_i X_{in}\right|^{\alpha} \leqq \left(\sum_{i=1}^{k} |t_i|^{\frac{\alpha}{\alpha-1}}\right)^{\alpha-1} \sum_{i=1}^{k} |X_{in}|^{\alpha} =: D_t \sum_{i=1}^{k} |X_{in}|^{\alpha},$$

and setting $D_t' = D_t / ||\mathbf{t}||^{\alpha}$, we obtain, for a suitable constant D, that

$$\sum_{n=1}^{\infty} \frac{1}{\hat{s}_n^{\alpha} (LL\hat{s}_n^2)^{\alpha/2}} \mathbf{E}|Y_n|^{\alpha} I\left[Y_n^2 > \varepsilon^2 \hat{s}_n^2 (LL\hat{s}_n^2)^{-1}\right]$$

$$\leqq \sum_{n=1}^{\infty} \frac{D_t}{\hat{s}_n^{\alpha} (LL\hat{s}_n^2)^{\alpha/2}} \sum_{i=1}^{k} \mathbf{E}|X_{in}|^{\alpha} I\left[\sum_{i=1}^{k} X_{in}^2 > \varepsilon^2 s_n^2 (LL\hat{s}_n^2)^{-1}\right]$$

$$\leqq kD_t' \sum_{n=1}^{\infty} \frac{1}{s_n^{\alpha} (LL\hat{s}_n^2)^{\alpha/2}} \mathbf{E}|X_{1n}|^{\alpha} I\left[\max_{1 \leqq i \leqq k} X_{in}^2 > \varepsilon^2 s_n^2 (kLL\hat{s}_n^2)^{-1}\right]$$

$$\leqq kD_t' \sum_{i=1}^{k} \sum_{n=1}^{\infty} \frac{1}{s_n^{\alpha} (LL\hat{s}_n^2)^{\alpha/2}} \mathbf{E}|X_{1n}|^{\alpha} I\left[X_{in}^2 > \varepsilon^2 s_n^2 (kLL\hat{s}_n^2)^{-1}\right]$$

$$\leqq kD_t' \Bigg(D + (k-1) \sum_{n=1}^{\infty} \frac{1}{s_n^{\alpha} (LLs_n^2 ||\mathbf{t}||^2)^{\alpha/2}} \mathbf{E}|X_{1n}|^{\alpha} \times$$

$$\mathbf{P}\left(X_{2n}^2 > \varepsilon^2 s_n^2 (kLLs_n^2 ||\mathbf{t}||^2)^{-1}\right) \Bigg) < \infty$$

for all $\varepsilon > 0$ via (2.3). Convergence of the last series is also implied by (2.3) since the fact $\text{cov}(f(Y), g(Y)) \geqq 0$ for f and g non-decreasing ensures

(2.5)

$$\mathbf{E}|X_{1n}|^{\alpha} I\left[|X_{1n}| > \varepsilon s_n (LLs_n^2)^{-1/2}\right] \geqq \mathbf{E}|X_{1n}|^{\alpha} \; \mathbf{P}\left(|X_{1n}| > \varepsilon s_n (LLs_n^2)^{-1/2}\right).$$

From here on, the argument is identical to that of Lemma 2 of Finkelstein [14]. □

REMARK 2.1. A sequence $\{X_{1n}, n \geqq 1\}$ of independent random variables with $\mathbf{E}X_{1n} = 0$, $\mathbf{E}X_{1n}^2 = \sigma_n^2 < \infty$, obeys the *two-sided LIL* if $S_{1n} = \sum_{j=1}^n X_{1n}$ and $s_n^2 = \sum_{j=1}^n \sigma_j^2$ are such that

$$\limsup_{n\to\infty} \pm(2s_n^2 LLs_n^2)^{-1/2} S_{1n} = 1 \quad \text{a.s.}. \tag{2.6}$$

The following lemma establishes Theorem 1.1, and extends this result to the non-i.i.d. case considered in Lemma 2.1.

LEMMA 2.2. *If $\mathcal{X}$ is t.i.i.d. with* $\mathbf{E}X = 0$, $\mathbf{E}X^2 = 1$, *or alternatively under the hypotheses of Lemma 2.1, the almost sure limit set of* $\{(2s_n^2 LLs_n^2)^{-1/2} M_k(n), n \geqq 1\}$ *is* $[-k^{-1/2}, 1]$.

PROOF. Via Lemma 2.1 or Lemma 2 of Finkelstein [14], the almost sure limit set $\mathcal{L}$ of the sequence $\{(2s_n^2 LLs_n^2)^{-1/2} \max_{1\leqq i\leqq k} S_{in},\ n \geqq 1\}$ is the image of the unit ball $\mathcal{B}_k$ of $\mathbf{R}^k$ under the mapping $(x_1, \dots, x_k) \to \max_{1\leqq i\leqq k} x_i$. Since $\mathcal{B}_k$ is convex and compact, $\mathcal{L}$ is an interval $[a, b]$ and $b \leq 1$ via $\sum_{i=1}^k x_i^2 \leqq 1$. In fact, $b = 1$ as follows by taking $x_1 = 1$, $x_i = 0$, $i \neq 1$. Likewise, the minimum value of $\max_{1\leqq i\leqq k} x_i$ in $\mathcal{B}_k$ is attained when $x_i < 0$, $1 \leqq i \leqq k$ and $\min_{1\leqq i\leqq k} |x_i|$ is a maximum. Since $k(\min_{1\leqq i\leqq k} |x_i|)^2 \leqq 1$ with equality for $x_i = -k^{-1/2}$, the conclusion follows. □

As in Section 1, let $S_{in} = \sum_{j=1}^n X_{ij}$ and $M_{k_n} = \max_{1\leqq i\leqq k_n} S_{in}$, where $\{k_n, n \geqq 1\}$ denotes a non-decreasing sequence of (possibly non-integer) positive numbers tending to ∞.

COROLLARY 2.1. *If either* (i) $\mathcal{X}$ *is t.i.i.d. with* $\mathbf{E}X = 0$, $\mathbf{E}X^2 = 1$, *or* (ii) $\mathcal{X}$ *satisfies the hypotheses of Lemma 2.1, then*

$$\liminf_{n\to\infty} (2s_n^2 LLs_n^2)^{-1/2} M_{k_n} \geqq 0 \ \textit{and}\ \limsup_{n\to\infty} (2s_n^2 LLs_n^2)^{-1/2} M_{k_n} \geqq 1 \ \textit{a.s.}. \tag{2.7}$$

PROOF. This follows from Lemma 2.2 and $M_{k_n} \geqq \max_{1\leqq i\leqq k} S_{in}$ for fixed $k \geqq 1$ and all large n. □

The next lemmas are geared to proving that all points between the lower bound $\liminf_{n\to\infty} b_n^{-1} M_{k_n}$ and the upper bound $\limsup_{n\to\infty} b_n^{-1} M_{k_n}$ are almost sure limit points of $b_n^{-1} M_{k_n}$, where, unless otherwise specified, $b_n = (2s_n^2 LLs_n^2)^{1/2}$. Let $\mathcal{X}$ be i.d., and, for any real r and s, define

$$H(r, s) = \mathbf{E}X^2 (L|X|)^r (LL|X|)^s. \tag{2.8}$$

As will be shown next, the finiteness of $H(r,s)$ for suitable choices of r and s plays an important role in the description of the limiting behaviour of $M_{k_n}(n)$.

LEMMA 2.3. *Let $k_n = n^\gamma (Ln)^{\beta_1}(LLn)^{\beta_2}$ where $\gamma \geqq 0$, $\beta_1 \geqq 0$, $\beta_1 \vee \beta_2 > 0$. If $\mathcal{X}$ is i.d. and $H(r,s) < \infty$ for some real r and s, then*

$$\lim_{n\to\infty} \left\{n^{1+\gamma}(Ln)^{\beta_1 - r}(LLn)^{\beta_2 - s}\right\}^{-1/2} \max_{1\leqq i \leqq k_n} |X_{in}| = 0 \quad a.s.. \tag{2.9}$$

Conversely, if $\mathcal{X}$ is t.i.i.d. and $H(r,s)=\infty$ for some real r and s, then

$$\limsup_{n\to\infty} \left\{n^{1+\gamma}(Ln)^{\beta_1 - r}(LLn)^{\beta_2 - s}\right\}^{-1/2} \max_{1\leqq i \leqq k_n} |X_{in}| = \infty \quad a.s.. \tag{2.10}$$

PROOF. Let $\lambda > 0$ and set $m_n = \{\lambda n^{1+\gamma}(Ln)^{\beta_1 - r}(LLn)^{\beta_2 - s}\}^{1/2}$. Now, if $\mathcal{X}$ is i.d.,

$$\begin{aligned}
&\sum_{n=1}^\infty \mathbf{P}\left(\max_{1\leqq i\leqq k_n} |X_{in}| \geqq m_n\right) \leqq \sum_{n=1}^\infty k_n \mathbf{P}\left(|X| \geqq m_n\right) \\
&= \sum_{n=1}^\infty k_n \sum_{j=n}^\infty \mathbf{P}\left(m_j \leqq |X| < m_{j+1}\right) \\
&\leqq \sum_{j=1}^\infty \mathbf{P}\left(m_j \leqq |X| < m_{j+1}\right) \sum_{n=1}^j n^\gamma (Ln)^{\beta_1}(LLn)^{\beta_2} \\
&= O\left(\sum_{j=1}^\infty j^{1+\gamma}(Lj)^{\beta_1}(LLj)^{\beta_2}\mathbf{P}\left(m_j \leqq |X| < m_{j+1}\right)\right) \\
&= O\left(\sum_{j=1}^\infty m_j^2 (Lm_j)^r (LLm_j)^s \mathbf{P}\left(m_j \leqq |X| < m_{j+1}\right)\right) = O(H(r,s)).
\end{aligned}$$

Thus, when $H(r,s) < \infty$, (2.9) follows from Borel–Cantelli lemma and the arbitrariness of $\lambda > 0$. When $\mathcal{X}$ is t.i.i.d., (2.10) follows similarly via

$$\begin{aligned}
\mathbf{P}\left(\max_{1\leqq i\leqq k_n} |X_{in}| \geqq m_n\right) &= 1 - (1 - \mathbf{P}(|X| \geqq m_n))^{[k_n]} \\
&\geqq 1 - \exp(-[k_n]\mathbf{P}(|X| \geqq m_n)) = (1+o(1))k_n \mathbf{P}(|X| \geqq m_n),
\end{aligned}$$

and the inequality, for some $C > 0$, $\infty = H(r,s) \leqq C\sum_{n=1}^\infty k_n \mathbf{P}(|X| \geqq m_n)$. □

The next proposition shows that when k_n is of *moderate size* (see Section 3), something beyond a second moment of X is needed to ensure the relative compactness of $(2nLLn)^{-1/2}M_{k_n}$ (see Remark 2.2 in the sequel).

PROPOSITION 2.1. *If $\mathcal{X}$ is t.i.i.d. with* $\mathbf{E}X = 0$, $\mathbf{E}X^2 = 1$ *and* $k_n = (Ln)^{\beta_1}(LLn)^{\beta_2}$ *where* $\beta_1 \geqq 0$ *and* $\beta_1 \vee \beta_2 > 0$, *a necessary condition for*

$$\limsup_{n\to\infty}(nLLn)^{-1/2} \max_{1\leqq i\leqq k_n} |S_{in}| < \infty \quad a.s. \tag{2.11}$$

is that

$$H(\beta_1, \beta_2 - 1) < \infty. \tag{2.12}$$

PROOF. If (2.12) fails to hold, by a straightforward modification of Lemma 2.3, taken with $\gamma = 0$, $r = \beta_1$ and $s = \beta_2 - 1$, we obtain that $\limsup_{n\to\infty} b_n^{-1} \max_{1\leqq i\leqq k_{n-1}} |X_{in}| = \infty$ a.s., where $b_n = (nLLn)^{1/2}$. Making use of the equality $\max_{1\leqq i\leqq k_{n-1}} |X_{in}| = \max_{1\leqq i\leqq k_{n-1}} |S_{in} - S_{in-1}|$, we see that we must then have a.s. either

(i) $\limsup_{n\to\infty} b_n^{-1} \max_{1\leqq i\leqq k_{n-1}} |S_{in}| = \infty$, or

(ii) $\limsup_{n\to\infty} b_n^{-1} \max_{1\leqq i\leqq k_{n-1}} |S_{in-1}| = \infty$.

Since either one of these two assertions implies the equality

$$\limsup_{n\to\infty} b_n^{-1} \max_{1\leqq i\leqq k_n} |S_{in}| = \infty,$$

we obtain the proposition in its contra-positive form. □

REMARK 2.2. The condition (2.11), $\limsup_{n\to\infty}(nLLn)^{-1/2} \max_{1\leqq i\leqq k_n} |S_{in}| < \infty$ a.s. implies

(i) $\limsup_{n\to\infty}(nLLn)^{1/2} \max_{1\leqq i\leqq k_n} S_{in} < \infty$ a.s., but the converse need not hold in general. This is, however, the case if the condition

(ii) $\limsup_{n\to\infty}(nLLn)^{-1/2} \max_{1\leqq i\leqq k_n} (-S_{in}) < \infty$ a.s. holds in addition to (i), that is, if we have a two-sided LIL version of (i). For symmetric random variables, (i) and (ii) are equivalent, since then $\mathcal{X}$ and $-\mathcal{X}$ are identically distributed.

The following lemma has interest in itself.

LEMMA 2.4. *Let* $\{b_n,\ n \geqq 1\}$ *be constants with*

$$0 < b_n \uparrow \infty, \quad \lim_{n\to\infty} b_{n+1}/b_n = 1,$$

and let $\{Z_n,\ n \geqq 1\}$ *be random variables satisfying* $b_{n+1}^{-1}(Z_{n+1} - Z_n) \to 0$ *a.s., and*

$$c = \liminf_{n\to\infty} b_n^{-1} Z_n < \limsup_{n\to\infty} b_n^{-1} Z_n = d \quad a.s., \tag{2.13}$$

where at least one of the constants c, d is finite. Then, the almost sure limit set of $b_n^{-1} Z_n$ is $[c, d]$.

PROOF. There is nothing to prove if $c = d$. Suppose that $c < d < \infty$. For any $y \in (c, d)$, with probability one, for infinitely many n

$$b_n^{-1} Z_n \geqq y > b_{n+1}^{-1} Z_{n+1}, \tag{2.14}$$

or equivalent $b_n^{-1} Z_n (1 - b_{n+1}^{-1} b_n) + b_{n+1}^{-1} (Z_n - Z_{n+1}) \geqq y - b_{n+1}^{-1} Z_{n+1} > 0$. Since the left side converges a.s. to zero, it follows with probability one that y is a limit point of $b_n^{-1} Z_n$. If rather $c > -\infty$, replace (2.14) by $b_n^{-1} Z_n \leqq y < b_{n+1}^{-1} Z_{n+1}$. □

LEMMA 2.5. *Let $\{b_n,\ n \geqq 1\}$ be constants with*

$$0 < b_n \uparrow \infty, \quad b_{n+1}/b_n \to 1,$$

and let $\mathcal{X}$ satisfy $0 \leqq c = \liminf_{n\to\infty} b_n^{-1} M_{k_n} < \limsup_{n\to\infty} b_n^{-1} M_{k_n} = d \leqq \infty$ a.s.. If

(i) *$k_{n+1} - k_n < 1$ for all large n and*

$$\max_{1 \leqq i \leqq k_n} b_n^{-1} |X_{in}| \to 0 \quad a.s., \tag{2.15}$$

or

(ii) *$c > 0$, (2.15) holds and*

$$\max_{k_n < i \leqq k_{n+1}} b_{n+1}^{-1} S_{in+1} \to 0 \quad a.s., \tag{2.16}$$

then the almost sure limit set of $b_n^{-1} M_{k_n}$ is $[c, d]$.

PROOF. Set $Z_n = M_{k_n}$. Now

$$Z_{n+1} - Z_n \geqq \max_{1 \leqq i \leqq k_n} (S_{in} + X_{in+1}) - M_{k_n} \geqq - \max_{1 \leqq i \leqq k_{n+1}} |X_{in+1}|$$

and, since under (i), $k_{n+1} - k_n < 1$ for all large n, we have

$$Z_{n+1} - Z_n = \max_{1 \leqq i \leqq k_n} (S_{in} - X_{in+1}) - M_{k_n} \leqq \max_{1 \leqq i \leqq k_{n+1}} |X_{in+1}|, \tag{2.17}$$

for all large n, whence the conclusion follows from Lemma 2.4. In Case (ii), with probability one, $M_{k_n} > \frac{c}{2} b_{n+1} > 0$ for all large n, whence (2.17) can be replaced by

$$\begin{aligned} b_{n+1}^{-1}(Z_{n+1} - Z_n) &\leqq b_{n+1}^{-1} \left(\left(M_{k_n} \vee \max_{k_n < i \leqq k_{n+1}} S_{in} \right) + \max_{1 \leqq i \leqq k_{n+1}} X_{i,n+1} - M_{k_n} \right) \\ &\leqq b_{n+1}^{-1} \max_{1 < i \leqq k_{n+1}} |X_{in+1}|, \end{aligned}$$

where we have used (2.16). □

LEMMA 2.6. *Let $s_n \to \infty$, and let $\mathcal{X}$ be c.i.d. and satisfy with probability one*

$$0 \leqq c = \liminf_{n\to\infty}(2s_n^2 LLs_n^2)^{-1/2}M_{k_n} < \limsup_{n\to\infty}(2s_n^2 LLs_n^2)^{-1/2}M_{k_n} = d \leqq \infty.$$

(i) *If, for all $\lambda > 0$,*

$$\sum_{n=1}^{\infty} k_n \mathbf{P}\left(|X_{1n}| > \lambda s_n (LLs_n^2)^{1/2}\right) < \infty, \tag{2.18}$$

and either

$$\frac{Lk_n}{LLs_n^2} \downarrow \beta_1 > 0, \qquad s_{n+1}^2 = s_n^2(1 + O(k_n^{-1})), \tag{ia}$$

or

$$\frac{Lk_n}{L_3 s_n^2} \downarrow \beta_2 > 0, \qquad s_{n+1}^2 = s_n^2(1 + o(1)), \tag{ib}$$

then the almost sure limit set of $(2s_n^2 LLs_n^2)^{-1/2}M_{k_n}$ is $[c,d]$.

(ii) *If $\mathcal{X}$ is i.d. with $H(\beta_1, \beta_2 - 1) < \infty$, $s_n^2 = n$ and $k_n = (Ln)^{\beta_1}(LLn)^{\beta_2}$, $\beta_1 \geqq 0$, $\beta_1 \vee \beta_2 > 0$, then the almost sure limit set of $(2s_n^2 LLs_n^2)^{-1/2}M_{k_n}$ is $[c,d]$.*

PROOF. Under (ia),

$$\frac{Lk_{n+1}}{Lk_n} \leqq \frac{LLs_{n+1}^2}{LLs_n^2},$$

implying that

$$L\left(\frac{k_{n+1}}{k_n}\right) < 2\beta_1 L\left(\frac{Ls_{n+1}^2}{Ls_n^2}\right) = O\left(\frac{1}{k_n Ls_n^2}\right)$$

for all large n. Thus

$$k_{n+1}/k_n < \exp\left\{O\left(\frac{1}{k_n Ls_n^2}\right)\right\}$$

whence

$$k_{n+1} - k_n < 2k_n O\left(\frac{1}{k_n Ls_n^2}\right) = o(1).$$

Under (ib), the latter relation becomes

$$k_{n+1} - k_n < 2k_n o\left(\frac{1}{(Ls_n^2)LLs_n^2}\right) = o(1).$$

Since for all $\lambda > 0$

$$\sum_{n=1}^{\infty} \mathbf{P}\left(\max_{1 \leqq i \leqq k_n} |X_{in}| > \lambda(s_n^2 LLs_n^2)^{1/2}\right) \leqq \sum_{n=1}^{\infty} k_n \mathbf{P}\left(|X_{1n}| > \lambda(s_n^2 LLs_n^2)^{1/2}\right) < \infty,$$

(2.15) holds with $b_n = (2s_n^2 LLs_n^2)^{1/2}$ and the conclusion follows from Lemma 2.5 (i).

In Case (ii), it is again true that $k_{n+1} - k_n < 1$ for all large n, so that taking $\gamma = 0$, $r = \beta_1$, $s = \beta_2 - 1$ in Lemma 2.3 and and choosing $b_n = (2nLLn)^{1/2}$, Lemma 2.5 (i) is again applicable. □

3. The upper limit

Throughout this section, the following assumptions and notation will be in force.

$\mathcal{X}$ is a totally independent, column-wise identically distributed array with

$$\text{(3.1)} \qquad \mathbf{E}(X_{1n}) = 0, \quad n \geqq 1, \quad s_n^2 = \sum_{j=1}^{n} \mathbf{E}(X_{1j}^2) \to \infty \quad \text{as} \quad n \to \infty.$$

Set

$$\text{(3.2)} \qquad S_{in} = \sum_{j=1}^{n} X_{ij} \quad \text{and} \quad M_{k_n} = \max_{1 \leqq i \leqq k_n} S_{in}.$$

The behaviour of M_{k_n} depends crucially upon the magnitude of k_n. Three main cases of interest may be delineated: *small* k_n, *moderate* k_n and *large* k_n. Moderate values of k_n are defined by

$$\text{(3.3)}' \qquad (Lk_n)/LLs_n^2 \to \beta_1 \in (0, \infty),$$

then k_n is “small” or “large” according as

$$\text{(3.3)}'' \qquad (Lk_n)/LLs_n^2 \to \beta_1 = 0 \text{ or } \infty.$$

At times, it is convenient to combine some moderate and small k_n by considering

$$\text{(3.4)} \qquad k_n = k_n(\beta_1, \beta_2) := (1 + o(1))(Ls_n^2)^{\beta_1}(LLS_n^2)^{\beta_2}, \ \beta_1 \geqq 0, \ \beta_1 \vee \beta_2 > 0.$$

Our first theorem covers moderate k_n as well as large values of k_n. For small k_n as in (3.3)″ with $\beta_1 = 0$, the classical LIL is recaptured (Corollary 3.1). For moderate k_n as in (3.3)′, the coefficient $2^{1/2}$ is transformed to $(2(1+\beta_1))^{1/2}$ (Theorem 3.2). On the other hand, for large values of k_n such as in (3.3)″ with $\beta_1 = \infty$, the order of magnitude of M_{k_n} itself is altered (Theorem 4.3). In the sequel, we use the convention $1/\infty = 0$.

THEOREM 3.1. *Let M_{k_n} be as in (3.2), where $\mathcal{X}$ satisfies (3.1). If, for all $\varepsilon > 0$,*

$$\sum_{n=1}^{\infty} k_n \mathbf{P}\left(|X_{1n}| > \varepsilon s_n (Lk_n)^{-1/2}\right) < \infty, \tag{3.5}$$

$$\sum_{j=1}^{n} \mathbf{E} X_{1j}^2 I\left[|X_{1j}| > \varepsilon s_j (Lk_j)^{-1/2}\right] = o(s_n^2) \quad \text{as } n \to \infty, \tag{3.6}$$

and

$$\text{(i) } s_n^2/Lk_n \uparrow \infty, \qquad \text{(ii) } (Lk_n)/LLs_n^2 \to \beta_1 \in (0,\infty], \tag{3.7}$$

then, with probability one,

$$\limsup_{n\to\infty} (2s_n^2 Lk_n)^{-1/2} M_{k_n} \leqq (1+\beta_1^{-1})^{1/2}. \tag{3.8}$$

Moreover, if either (iii) $\beta_1 < \infty$ *or* (iv) $\beta_1 = \infty$, $Lk_n = c(L_q s_n^2)^{\theta}$ *for* $c > 0$ *and, either* $q = 1$, $\theta > 0$ *or* $q = 2$, $\theta > 1$, *then equality holds in* (3.8).

PROOF. There exists a positive sequence $\{\bar{\varepsilon}_n,\ n \geq 1\}$ with $\bar{\varepsilon}_n \downarrow 0$ such that (3.5) and (3.6) hold with ε replaced by $\bar{\varepsilon}_n$ and $\bar{\varepsilon}_j$, respectively (see e.g. Teicher [27]). Denote respectively these strengthened versions by (3.5)* and (3.6)*. Set

$$d_j = \bar{\varepsilon}_j s_j (Lk_j)^{-1/2}, \qquad d_j' = s_j (Lk_j)^{-1/2}. \tag{3.9}$$

Since $d_n/d_n' = \bar{\varepsilon}_n = o(1)$ and via (i), $0 < d_n' \uparrow \infty$, it follows that $\max_{1\leqq j\leqq n} d_j = o(d_n')$, so that d_n may be supposed non-decreasing. Define

$$X_{ij} = (X_{ij} I[|X_{ij}| \leqq d_j] - \mathbf{E} X_{1j} I[|X_{1j}| \leqq d_j]) + X_{ij} I[|X_{ij}| > d_j] - \mathbf{E} X_{1j} I[|X_{1j}| > d_j] =: X_{ij}' + X_{ij}'' - \rho_j. \tag{3.10}$$

Now, by (3.5)*,

$$\sum_{n=1}^{\infty} \mathbf{P}\left(\bigcup_{1\leqq i\leqq k_n} \{|X_{in}| > d_n\}\right) \leqq \sum_{n=1}^{\infty} k_n \mathbf{P}(|X_{1n}| > d_n) < \infty,$$

so that, by the Borel–Cantelli lemma, $\mathbf{P}\left(\max_{1\leqq i\leqq k_n} |X_{in}| > d_n \text{ i.o.}\right) = 0$, implying for $b_n = s_n(2Lk_n)^{1/2}$ that

$$b_n^{-1} \max_{1\leqq i\leqq k_n} \sum_{j=1}^{n} |X_{ij}''| \to 0 \quad \text{a.s.}. \tag{3.11}$$

Moreover, by (3.9),

$$\begin{aligned}\left|\sum_{j=1}^{n}\rho_j\right| &\leqq \sum_{j=1}^{n}\mathbf{E}|X_{1j}|\left(I\left[d_j<|X_{1j}|\leqq d_n'\right]+I\left[|X_{1j}|>d_n'\right]\right)\\ &\leqq d_n'\sum_{j=1}^{n}\mathbf{P}\left(|X_{1j}|>d_j\right)+\frac{1}{d_n'}\sum_{j=1}^{n}\mathbf{E}X_{1j}^2I\left[|X_{1j}|>d_n'\right],\end{aligned}$$

implying via (3.5)* and (3.6)* that

$$\sum_{j=1}^{n}\rho_j=o(b_n). \tag{3.12}$$

Furthermore, since $\mathbf{E}X_{1j}=0$, for $j\geqq 1$,

$$\begin{aligned}\mathbf{E}X_{1j}^2-\mathbf{E}(X_{1j}')^2&=\mathbf{E}X_{1j}^2I\left[|X_{1j}|>d_j\right]+\mathbf{E}^2X_{1j}I\left[|X_{1j}|>d_j\right]\\ &\leqq 2\mathbf{E}X_{1j}^2I\left[|X_{1j}|>d_j\right].\end{aligned}$$

By (3.6)*, setting ${s_n'}^2=\sum_{j=1}^{n}\mathbf{E}{X_{1j}'}^2$, this implies that

$$s_n^2-{s_n'}^2\leqq 2\sum_{j=1}^{n}\mathbf{E}X_{1j}^2I\left[|X_{1j}|>d_j\right]=o(s_n^2)\text{ as }n\to\infty,$$

so that $s_n=(1+o(1))s_n'$. Define $S_{in}'=\sum_{j=1}^{n}X_{ij}'$ and $M_{k_n}'=\max_{1\leqq i\leqq k_n}S_{in}'$. Observe from (3.6)* that

$$\begin{aligned}\mathbf{E}X_{1n}^2&=\mathbf{E}X_{1n}^2I\left[|X_{1n}|\leqq\bar{\varepsilon}_ns_n(Lk_n)^{-1/2}\right]+\mathbf{E}X_{1n}^2I\left[|X_{1n}|>\bar{\varepsilon}_ns_n(Lk_n)^{-1/2}\right]\\ &\leqq\bar{\varepsilon}_n^2s_n^2(Lk_n)^{-1}+\sum_{j=1}^{n}\mathbf{E}X_{1j}^2I\left[|X_{1j}|>\bar{\varepsilon}_js_j(Lk_j)^{-1/2}\right]=o(s_n^2).\end{aligned}$$

Therefore, $s_{n+1}/s_n=(1+o(1))s_{n+1}'/s_n'\to 1$, and, for any $\delta>1$, there exists an integer sequence $\{n_r,\ r\geqq r_0\}$ such that $(s_{n_r}')^2\leqq\delta^r<(s_{n_{r+1}}')^2$ for $r\geqq r_0$. Then $(s_{n_r}')^2=(1+o(1))\delta^r$ as $r\to\infty$. Let $\chi_n=(Lk_n)^{1/2}$. For any fixed $\lambda > \mu > 1$, we have for all large r

$$\begin{aligned} p_r(\lambda) := \mathbf{P}\left(\max_{n_{r-1}<n\leqq n_r}\ \max_{1\leqq i\leqq k_{n_r}} S'_{in} > \lambda^{1/2} b_{n_r}\right) \\ = \mathbf{P}\left(\max_{1\leqq i\leqq k_{n_r}}\ \max_{n_{r-1}<n\leqq n_r} S'_{in} > \lambda^{1/2} b_{n_r}\right) \\ \leqq k_{n_r}\mathbf{P}\left(\max_{1\leqq n\leqq n_r} S'_{1n} > s_{n_r}(2\lambda L k_{n_r})^{1/2}\right) \\ \leqq k_{n_r}\mathbf{P}\left(\max_{1\leqq n\leqq n_r} S'_{1n} > s'_{n_r}(2\mu L k_{n_r})^{1/2}\right) \\ = k_{n_r}\mathbf{P}\left(\max_{1\leqq n\leqq n_r} S'_{1n} > (2\mu)^{1/2}\chi_{n_r} s'_{n_r}\right) \end{aligned} \tag{3.13}$$

where we have used the fact that $s_n/s'_n \to 1$. By (3.9) and (3.10), for all large n and $1 \leqq j \leqq n$,

$$|X'_{ij}| \leqq 2d_j \leqq 2d_n = 2\bar{\varepsilon}_n s_n (Lk_n)^{-1/2} \leqq \{3\bar{\varepsilon}_n (Lk_n)^{-1/2}\} s'_n =: c_n s'_n. \tag{3.14}$$

We now make use of Lemma 10.2.1, p. 350 of Chow and Teicher [2], cited for convenience in Fact 1 below.

FACT 1. *Let $h(x) = (1+x)\log(1+x) - x$ for $x > -1$. Let $\{\xi_n, n \geqq 1\}$ be independent random variables with $\mathbf{E}\xi_n = 0$ and $0 < \bar{s}_n^2 = \sum_{j=1}^n \mathbf{E}\xi_j^2 < \infty$, and set $\zeta_n = \sum_{j=1}^n \xi_j$, $n \geqq 1$. Then, if c_n, x_n, a_n and λ_n are positive numbers such that $c_n x_n \leqq a_n$ and $\mathbf{P}(\xi_j \leqq c_n s_n) = 1$ for $1 \leqq j \leqq n$, we have*

$$\mathbf{P}\left(\max_{1\leqq j\leqq n} \zeta_j \geqq \lambda_n x_n \bar{s}_n\right) \leqq \exp(-a_n^{-2} h(\lambda_n a_n) x_n^2). \tag{3.15}$$

We apply Fact 1, taken with $\xi_j = X'_{ij}$, $\zeta_n = S'_{in}$, $\bar{s}_n = s'_n$, $x_n = \chi_n$, $\lambda_n = (2\mu)^{1/2}$, $a_n = 3\bar{\varepsilon}_n$, $c_n = \{3\bar{\varepsilon}_n(Lk_n)^{-1/2}\}$ as in (3.14), and $n = n_r$. Since $\bar{\varepsilon}_n \to 0$, we see that $a_n = c_n x_n = 3\bar{\varepsilon}_n =: \varepsilon_n \to 0$. Therefore, by (3.13), (3.15) and making use of the observation that $h(x)/x^2 \to 1/2$ as $x \to 0$, we obtain that, for any fixed $\lambda > \mu > \nu > 1$ and all large r,

$$\begin{aligned} p_r(\lambda) \leqq k_{n_r}\exp\left(-\varepsilon_{n_r}^2 h((2\mu)^{1/2}\varepsilon_{n_r}) L k_{n_r}\right) \leqq \exp(-(\nu-1) L k_{n_r}) \\ =: \exp(-(\nu-1) u_{n_r} LL s_{n_r}^2 \leqq (r\log\delta)^{-(\nu-1)u_{n_r}}, \end{aligned} \tag{3.16}$$

where, by (3.7) (ii), $u_n = (Lk_n)/LLs_n^2 \to \beta_1$ as $n \to \infty$.

By (3.16), a proper choice of $\nu \in (1, \lambda)$ ensures that $\sum_r p_r(\lambda) < \infty$ if either

$$\beta_1 = \infty \quad \text{and} \quad \lambda > 1, \tag{3.17$'$}$$

(in which case any choice of $\nu \in (1, \lambda)$ will do) or

$$(3.17)'' \qquad 0 < \beta_1 < \infty \quad \text{and} \quad \lambda > 1 + \beta_1^{-1},$$

(in which case any choice of $\nu \in ((1 + \beta_1^{-1}), \lambda)$ will do).

Since, by (3.7) (i), $s_n^2/Lk_n \uparrow$, and recalling $b_n = (2s_n^2 Lk_n)^{1/2}$, we have

$$\begin{aligned} b_{n_{r-1}}/b_{n_r} &= (s^2_{n_{r-1}}/s^2_{n_r})\{s_{n_r}/(Lk_{n_r})^{1/2}\}/\{s_{n_{r-1}}/(Lk_{n_{r-1}})^{1/2}\} \\ &\geqq s^2_{n_{r-1}}/s^2_{n_r} \to \delta^{-1} \text{ as } r \to \infty. \end{aligned}$$

Thus, for any fixed $\rho > \lambda > 1$ if $\beta_1 = \infty$ (resp. $\rho > \lambda > 1 + \beta_1^{-1}$ if $0 < \beta_1 < < \infty$), we obtain that

$$\begin{aligned} &\mathbf{P}\left(M'_{k_n} > \rho^{1/2}\delta b_n \text{ i.o.}\right) \\ (3.18) \qquad &\leqq \mathbf{P}\left(\max_{n_{r-1}<n\leqq n_r} \max_{1\leqq i\leqq k_{n_r}} S'_{in} > \rho^{1/2}\delta b_{n_{r-1}} \text{ i.o.}\right) \\ &\leqq \mathbf{P}\left(\max_{n_{r-1}<n\leqq n_r} \max_{1\leqq i\leqq k_{n_r}} S'_{in} > \lambda^{1/2} b_{n_r} \text{ i.o.}\right) = 0, \end{aligned}$$

via (3.13), (3.17) and the Borel–Cantelli lemma. Hence, by letting δ decrease to one and ρ decrease to either 1 or $1 + \beta_1^{-1}$, we obtain that, for $\beta_1 = \infty$,

$$(3.19)' \qquad \limsup_{n\to\infty} b_n^{-1} M'_{k_n} \leqq 1 \quad \text{a.s.},$$

whereas, for $0 < \beta_1 < \infty$,

$$(3.19)'' \qquad \limsup_{n\to\infty} b_n^{-1} M'_{k_n} \leqq (1 + \beta_1^{-1})^{1/2} \quad \text{a.s.},$$

Then, (3.8) follows by combining (3.11), (3.12) and (3.19).

To obtain the reverse inequalites in (3.19)$'$ and (3.19)$''$, define, for $\gamma \in \in (0, 1)$, $\delta > 1$ and $\alpha > 0$,

$$D_{ir} = \left\{S'_{in_r} - S'_{in_{r-1}} > \alpha^{1/2}(1-\gamma)^2(1-\delta^{-1})^{1/2} b_{n_r}\right\}, \quad A_r = \bigcup_{1\leqq i\leqq k'_r} D_{ir},$$

where $k'_r := k_{n_{r-1}}$ and note that

$$\begin{aligned} \mathbf{P}(D_{ir}) &= \mathbf{P}\left(S'_{in_r} - S'_{in_{r-1}} > (1-\gamma)^2[(1-\delta^{-1})2\alpha s^2_{n_r} Lk_{n_r}]^{1/2}\right) \\ &= \mathbf{P}\left(S'_{in_r} - S'_{in_{r-1}} > (1-\gamma)^2 g_r x_r\right), \end{aligned}$$

where

$$g_r^2 = (s'_{n_r})^2 - (s'_{n_{r-1}})^2 = (1+o(1))\,(s'_{n_r})^2\,(1-\delta^{-1}) = (1+o(1)) s^2_{n_r}(1-\delta^{-1})$$

and

$$x_r^2 = (1+o(1)) 2\alpha Lk_{n_r}.$$

We will make use of Lemma 2 of Teicher [28], or Lemma 10.2.2, p. 353 of Chow and Teicher [2], which we cite for convenience in Fact 2 below.

FACT 2. *Let $\{\xi_n,\ n \geqq 1\}$ be independent random variables with $\mathbf{E}\xi_n = 0$ and $0 < \overline{s}_n^2 = \sum_{j=1}^n \mathbf{E}\xi_j^2 < \infty$, and set $\zeta_n = \sum_{j=1}^n \xi_j$, $n \geqq 1$. Then, if d_n and $\overline{x}_n$ are positive numbers such that $\overline{x}_n \to \infty$, $d_n\overline{x}_n/\overline{s}_n \to 0$ and $\mathbf{P}(|\xi_j| \leqq d_n) = 1$ for $1 \leqq j \leqq n$, for every $\gamma \in (0,1)$, there exists a $C_\gamma \in \left(\frac{1}{4}, \frac{1}{2}\right)$ such that for all large n*

$$\mathbf{P}\left(\zeta_n > (1-\gamma)^2\overline{s}_n\overline{x}_n\right) \geqq C_\gamma \exp\left(-\frac{1}{2}(1-\gamma)(1-\gamma^2)\overline{x}_n^2\right). \tag{3.20}$$

Since $|X'_{ij}| \leqq 3\overline{\varepsilon}_n s'_n (Lk_n)^{-1/2} =: d_n$ for $1 \leqq j \leqq n$, and $d_{n_r}x_r/g_r \to 0$ as $r \to \infty$, Fact 2, taken with $\xi_j = X'_{in_{r-1}+j}$, $\zeta_{in} = S'_{in_r} - S'_{in_{r-1}}$, $\overline{s}_n = g_r$, $\overline{x}_n = x_r$ and $n = \nu_r := n_r - n_{r-1}$, implies that, for some $C_\gamma \in \left(\frac{1}{4}, \frac{1}{2}\right)$ and all large r,

$$\mathbf{P}(D_{1r}) \geqq C_\gamma \exp(-(1-\gamma)(1-\gamma^2)x_r^2/2) \geqq C_\gamma \exp(-\alpha(1-\gamma)Lk_{n_r}).$$

Now either, via (iii),

$$Lk_{n_r} = (1+o(1))\beta_1 LLs_n^2 = (1+o(1))\beta_1 L(rL\delta), \quad 0 < \beta_1 < \infty,$$

or, via (iv),

$$Lk_{n_r} = (1+o(1))c(L_{q-1}(rL\delta))^\theta.$$

Thus in the either case, $Lk_{n_{r-1}} = (1+o(1))Lk_{n_r}$ whence, with $k'_r = k_{n_{r-1}}$,

$$\begin{aligned}\mathbf{P}(A_r) = 1 - (1-\mathbf{P}(D_{1r}))^{[k'_r]} &\geqq 1 - (1 - C_\gamma\exp(-\alpha(1-\gamma)Lk_{n_r}))^{[k'_r]}\\ &\geqq 1 - \exp(-[k'_r]C_\gamma\exp(-\alpha(1-\gamma)Lk_{n_r}))\\ &= 1 - \exp(-C_\gamma\exp((1-\alpha(1-\gamma))(1+o(1))Lk_{n_r})).\end{aligned}$$

Choose in the sequel $\alpha = 1 + \beta_1^{-1}$ if (iii) $0 < \beta_1 < \infty$ or $\alpha = 1$ if (iv) $\beta_1 = \infty$. In Case (iii),

$$\begin{aligned}\mathbf{P}(A_r) &= 1 - \exp(-C_\gamma\exp((\gamma(\beta_1+1)-1)(1+o(1))LLs_{n_r}^2))\\ &\geqq 1 - \exp\left(-C_\gamma(rL\delta)^{-1+\frac{\gamma}{2}(\beta_1+1)}\right) = (1+o(1))C_\gamma(rL\delta)^{-1+\frac{\gamma}{2}(\beta_1+1)},\end{aligned}$$

whence $\sum_r \mathbf{P}(A_r) = \infty$ for all sufficiently small $\gamma > 0$. In Case (iv) with $\alpha = 1$, the series again diverges since $\mathbf{P}(A_r) \geqq 1 - \exp(-C_\gamma\exp(\gamma u_{n_r}LLs_{n_r}^2)) \to 1$ in view of $u_n \to \infty$. Consequently, the Borel–Cantelli lemma implies in either case that

$$\begin{aligned}\mathbf{P}\left(\max_{1\leqq i\leqq k_{n_{r-1}}}(S'_{in_r} - S'_{in_{r-1}}) > \alpha^{1/2}(1-\gamma)^2(1-\delta^{-1})^{1/2}b_{n_r} \text{ i.o.}\right)&\\ = \mathbf{P}(A_r \text{ i.o.}) = 1.&\end{aligned} \tag{3.21}$$

Select $\delta > 1$ so large that

$$(1-\gamma)^2(1-\delta^{-1})^{1/2} - 4\delta^{-1/2} > (1-\gamma)^3,$$

implying for all large r that

$$2\delta^{-1/2}b_{n_r} = (1+o(1))2b_{n_{r-1}} > b_{n_{r-1}}$$

and

$$\alpha^{1/2}(1-\gamma)^2(1-\delta^{-1})^{1/2}b_{n_r} - 2\alpha^{1/2}b_{n_{r-1}} > \alpha^{1/2}(1-\gamma)^3 b_{n_r}.$$

If

$$B_r := \left\{ \min_{1 \leqq i \leqq k_{n_{r-1}}} S_{in_{r-1}} > -2\alpha^{1/2}b_{n_{r-1}} \right\},$$

by replacing X_{ij} by $-X_{ij}$ in (3.19)′ and (3.19)″ we get

$$\mathbf{P}(B_r \text{ holds for all large } r) = 1. \tag{3.22}$$

Thus, via (3.21) and (3.22),

$$\mathbf{P}\left(\max_{1 \leqq i \leqq k_{n_{r-1}}} S'_{in_r} > \alpha^{1/2}(1-\gamma)^3 b_{n_r} \text{ i.o.} \right) \geqq \mathbf{P}(A_r B_r \text{ i.o.}) = 1,$$

so that, a fortiori, with probability one,

$$\limsup_{n\to\infty} b_n^{-1} M'_{k_n} \geqq \limsup_{n_r\to\infty} b_{n_r}^{-1} M'_{k_{n_r}} \geqq \alpha^{1/2}(1-\gamma)^3,$$

where the right-hand side increases to $\alpha^{1/2}$ as γ decreases to zero. In conjunction with (3.11), (3.12), this yields equality in (3.8). This completes the proof of Theorem 3.1. □

In Section 4, it will be shown that if $s_n^2 \geqq cn^\alpha$ for some positive c, α, and k_n satisfies (3.7) with $\beta_1 = \infty$, then (3.5), (3.6) imply $(2s_n^2 Lk_n)^{-1/2}M_{k_n} \to 1$ a.s. .

The next theorem will be used in Section 4 for the proof of Theorem 4.2.

THEOREM 3.2. *Let $\mathcal{X}$ satisfy* (3.1) *and let k_n be governed by* (3.3)′. *If, for all $\varepsilon > 0$,*

$$\sum_{n=1}^{\infty} k_n \mathbf{P}\left(|X_{1n}| > \varepsilon s_n (LLs_n^2)^{-1/2}\right) < \infty, \tag{3.23}$$

and

$$\sum_{j=1}^{n} \mathbf{E}X_{1j}^2 I\left[|X_{1j}| > \varepsilon s_j (LLs_j^2)^{-1/2}\right] = o(s_n^2) \quad \text{as } n \to \infty, \tag{3.24}$$

then

$$\limsup_{n\to\infty}(2s_n^2 LLs_n^2)^{-1/2}M_{k_n} = (1+\beta_1)^{1/2} \quad a.s.. \tag{3.25}$$

PROOF. Since $Lk_n = (1+o(1))\beta_1 LLs_n^2$, the result follows directly from Theorem 3.1. □

COROLLARY 3.1. *Let $\mathcal{X}$ satisfy conditions* (3.23), (3.24) *of Theorem 3.2. Let $1 \leqq k_n' \leqq k_n(0,\beta_2)$ for some $\beta_2 > 0$, where $k_n(0,\beta_2)$ is governed by* (3.4). *Then*

$$\limsup_{n\to\infty}(2s_n^2 LLs_n^2)^{-1/2}M_{k_n'} = 1 \quad a.s.. \tag{3.26}$$

PROOF. It is easy to see (see Teicher [27]) that the classical LIL (that is, (3.26) with $k_n' = 1$) holds under (3.23), (3.24). Since

$$M_1(n) \leqq M_{k_n'}(n) \leqq M_{k_n(0,\beta_2)}(n) \leqq M_{k_n(\beta_1,\beta_2)}(n)$$

for arbitrary $\beta_1 > 0$, the conclusion follows from Theorem 3.2 as $\beta_1 \downarrow 0$. □

In what follows, a product $\prod_{i=k}^{m}$ with $m < k$ is to be interpreted as unity. Our next corollary replaces the conditions of Theorem 3.2 by (3.27) at the cost of a small decrease in k_n. Define

$$\mathcal{H}(r,s) = \sup_{n\geqq 1} \mathbf{E}X_{1n}^2 (L|X_{1n}|)^r (LL|X_{1n}|)^s.$$

COROLLARY 3.2. *Let $\mathcal{X}$ satisfy* (3.1) *and* $\inf_{n\geqq 1}\mathbf{E}X_{1n}^2 = \sigma^2 > 0$. *If*

$$\mathcal{H}(1+\beta_1, 2+\beta_2) < \infty \text{ where } \beta_1 \geqq 0,\ \beta_1 \vee \beta_2 > 0, \tag{3.27}$$

$$\overline{k}_n = \overline{k}_n(\beta_1,\beta_2) = \frac{(Ls_n^2)^{\beta_1}(LLs_n^2)^{\beta_2}}{(L_{m+1}s_n^2)^{1+\eta}\prod_{i=3}^{m} L_i s_n^2}, \tag{3.28}$$

for some $m \geqq 2$ and $\eta > 0$, then

$$\limsup_{n\to\infty}(2s_n^2 LLs_n^2)^{-1/2}M_{\overline{k}_n} = (1+\beta_1)^{1/2} \quad a.s.. \tag{3.29}$$

PROOF. It suffices to verify (3.23), (3.24) of Theorem 3.2. Set $d_n'' = \varepsilon s_n (LLs_n^2)^{-1/2}$. Now, via (3.27), nothing that $\mathbf{E}X_{1n}^2$ is bounded away from zero and infinity, we obtain that, for suitable constants C and C'

$$\overline{k}_n \mathbf{P}\left(|X_{1n}| > d_n''\right) \leqq \frac{C\overline{k}_n}{\varepsilon^2 s_n^2 (Ls_n^2)^{1+\beta_1}(LLs_n^2)^{1+\beta_2}} \leqq \frac{C'}{n(L_{m+1}n)^{1+\eta}\prod_{i=1}^{m} L_i n},$$

which is summable. Moreover, again employing (3.27), we obtain that

$$\mathbf{E}X_{1j}^2 I\left[|X_{1j}| > d_j''\right] \leqq \frac{C}{(Ls_j^2)^{1+\beta_1}(LLs_j^2)^{1+\beta_2}} \leqq \frac{C'}{(Lj)^{1+\beta_1}(LLj)^{1+\beta_2}},$$

implying $\sum_{j=1}^n \mathbf{E}X_{ij}^2 I\left[|X_{1j}| > d_j''\right] = o(n) = o(s_n^2)$, which is (3.24). The conclusion follows from Theorem 3.2 or Corollary 3.1 according as $\beta_1 > 0$ or $\beta_1 = 0 < \beta_2$. □

COROLLARY 3.3. *Let $\mathcal{X}$ satisfy* (3.1) *and* $\inf_{n \geqq 1} \mathbf{E}X_{1n}^2 = \sigma^2 > 0$. *If*

$$k_n = k_n(\beta_1, \beta_2) = (1 + o(1))(Ls_n^2)^{\beta_1}(LLs_n^2)^{\beta_2}, \quad \beta_1 \geqq 0, \ \beta_1 \vee \beta_2 > 0,$$

and, for some $\delta > 0$,

$$\begin{aligned} \mathcal{H}(1 + \beta_1 + \delta, \beta_2 \wedge 0) &< \infty \quad \text{if } \beta_1 > 0, \\ \mathcal{H}(1, 2 + \beta_2 + \delta) &< \infty \quad \text{if } \beta_1 = 0 < \beta_2, \end{aligned} \tag{3.30}$$

then

$$\limsup_{n \to \infty} (2s_n^2 LLs_n^2)^{-1/2} M_{k_n} = (1 + \beta_1)^{1/2} \quad a.s.. \tag{3.31}$$

PROOF. If $\beta_1 > 0$ and $0 < \delta' < \delta$, (3.30) ensures $\mathcal{H}(1+\beta_1+\delta', \beta_2+2) < \infty$. Since

$$\overline{k}_n(\beta_1, \beta_2) \leqq k_n(\beta_1, \beta_2) \leqq \overline{k}_n(\beta_1 + \delta', \beta_2)$$

for all large n, Corollary 3.2 guarantees that

$$(1 + \beta_1)^{1/2} \leqq \limsup_{n \to \infty} (2s_n^2 LLs_n^2)^{-1/2} M_{k_n} = (1 + \beta_1 + \delta')^{1/2} \quad \text{a.s.},$$

and (3.31) follows as $\delta' \downarrow 0$. The argument is analogous when $\beta_1 = 0$. □

4. The lower limit

In this section, the behaviour of the lower limit of $b_n^{-1} M_{k_n}$ is addressed. For moderate k_n, Theorem 4.2 specifies the lower limit while, for large k_n, Theorem 4.3 asserts the equality of the upper and lower limits.

THEOREM 4.1. *Let $\mathcal{X}$ be t.i. and c.i.d. with*

$$\mathbf{E}X_{1n} = 0 \quad \text{and} \quad s_n^2 = \sum_{j=1}^n \mathbf{E}X_{ij}^2 \to \infty.$$

If

$$(Lk_n)/(LLs_n^2) \to \beta_1 \in [0, \infty) \quad \text{as } n \to \infty, \tag{4.1}$$

and, for each $\varepsilon > 0$.

$$\sum_{j=1}^{n} \mathbf{E}X_{ij}^2 I\left[|X_{ij}| > \varepsilon s_j (Lk_j)^{-1/2}\right] = o\left(\frac{s_n^2 LLs_n^2}{k_n}\right), \tag{4.2}$$

then

$$\lim_{n\to\infty} (2s_n^2 LLs_n^2)^{-1/2} M_{k_n} = \beta_1^{1/2} \quad \textit{in probability} \tag{4.3}$$

and

$$\liminf_{n\to\infty} (2s_n^2 LLs_n^2)^{-1/2} M_{k_n} \leqq \beta_1^{1/2} \quad \textit{a.s.}. \tag{4.4}$$

PROOF. Since (4.4) is an immediate consequence of (4.3), it suffices to prove the latter. Suppose initially that $\beta_1 > 0$. Then, as in choosing $\bar{\varepsilon}_n$ in the proof of Theorem 3.1, there exists a sequence $\bar{\varepsilon}_n \downarrow 0$ such that (4.2) holds with ε replaced by $\bar{\varepsilon}_n$. Denote by (4.2)* this strengthened version of (4.2). Let X'_{ij}, S'_{ij}, and s'_n be as in the proof of Theorem 3.1 (see e.g. (3.10)–(3.13)), but with $d_n = \bar{\varepsilon}_n s_n (LLs_n^2)^{-1/2}$. Let $b_n = (2s_n^2 LLs_n^2)^{1/2}$ and define, for $n \geqq 1$,

$$T_{in} = \sum_{j=1}^{n} (X_{ij} - X'_{ij}) \quad \text{and} \quad M'_{k_n} = \max_{1\leqq i\leqq k_n} S'_{in}. \tag{4.5}$$

In view of

$$M'_{k_n} - \max_{1\leqq i\leqq k_n} |T_{in}| \leqq M_{k_n} \leqq M'_{k_n} + \max_{1\leqq i\leqq k_n} |T_{in}|, \tag{4.6}$$

it suffices to verify that

$$b_n^{-1} M'_{k_n} \to \beta_1^{1/2} \quad \text{in probability}, \tag{4.7}$$

and

$$b_n^{-1} \max_{1\leqq i\leqq k_n} |T_{in}| \to 0 \quad \text{in probability}. \tag{4.8}$$

Notice that (4.2) implies (3.6), which in turn, as in the proof of Theorem 3.1, implies $s_n = (1 + o(1))s'_n$. Therefore, via Fact 2, we see that for any $\gamma \in (0,1)$ and all large n,

$$\begin{aligned}
&\mathbf{P}\left(M'_{k_n} \leqq (1-\gamma)^2 \beta_1^{1/2} b_n\right) \\
&= \left\{1 - \mathbf{P}\left(S'_{1n} > (1-\gamma)^2 s'_n (2\beta_1(1+o(1))LLs_n^2)^{1/2}\right)\right\}^{[k_n]} \\
&\leqq \left\{1 - C_\gamma \exp\left(-(1-\gamma)(1-\gamma^2)\beta_1(1+o(1))LLs_n^2\right)\right\}^{[k_n]} \\
&\leqq \exp\left(-[k_n] C_\gamma e^{-(1-\gamma)\beta_1 LLs_n^2}\right) \\
&\leqq \exp\left(-C_\gamma e^{(\gamma\beta_1 + o(1))LLs_n^2}\right) \leqq \exp\left(-C_\gamma (Ls_n^2)^{\gamma\beta_1/2}\right) = o(1).
\end{aligned} \tag{4.9}$$

On the other hand, for any $\lambda > 1$,

$$\begin{aligned}\mathbf{P}\Big(M'_{k_n} > (\lambda\beta_1)^{1/2}b_n\Big) &= 1 - \Big\{1 - \mathbf{P}\Big(S'_{1n} > (\lambda\beta_1)^{1/2}b_n\Big)\Big\}^{[k_n]} \\ &\leqq [k_n]\mathbf{P}\Big(S'_{1n} > (\lambda\beta_1)^{1/2}b_n\Big) = o(1),\end{aligned} \tag{4.10}$$

since via the inequality employed in (3.16), if $1 < \mu < \rho < \lambda$,

$$\begin{aligned}k_n\mathbf{P}\Big(S'_{1n} > (\lambda\beta_1)^{1/2}b_n\Big) &= k_n\mathbf{P}\Big(S'_{1n} > (\lambda\beta_1)^{1/2}(1+o(1))s'_n(2LLs_n^2)^{1/2}\Big) \\ \leqq k_n \exp\Big(-\varepsilon_n^{-2}h(\varepsilon_n(2\rho\beta_1)^{1/2})LLs_n^2\Big) &\leqq \exp\big((\beta_1 - \rho\beta_1 + o(1))LLs_n^2\big) \\ &\leqq \exp\big(-(\mu-1)\beta_1 LLs_n^2\big) = o(1).\end{aligned}$$

Clearly, (4.9) and (4.10) imply (4.7). Furthermore, for any $\beta_1 > 0$,

$$\begin{aligned}\mathbf{P}\Big(\max_{1\leqq i\leqq k_n} |T_{in}| > \delta b_n\Big) &\leqq k_n\mathbf{P}(|T_{1n}| > \delta b_n) \\ &\leqq \frac{k_n}{\delta^2 b_n^2}\sum_{j=1}^{n}\mathbf{E}X_{1j}^2 I\left[|X_{1j}| > d_j\right] = o(1),\end{aligned}$$

according to (4.2)*, yielding (4.8) and hence (4.3) when $\beta_1 > 0$.

Suppose next that $(Lk_n)/LLs_n^2 = o(1)$, or equivalently that $k_n/(Ls_n^2)^\beta \to \to 0$ for all $\beta > 0$. Fix an arbitrary $\varepsilon > 0$, and set $k'_n = (Ls_n^2)^\beta$, for $0 < \beta < \varepsilon^2/3$. Then, for all large n,

$$\mathbf{P}\Big(M_{k_n} > \varepsilon s_n(LLs_n^2)^{1/2}\Big) \leqq \mathbf{P}\Big(M_{k'_n} > (3\beta s_n^2 LLs_n^2)^{1/2}\Big) = o(1), \tag{4.11}$$

via the version of (4.3) obtained by replacing k_n by k'_n. Also, by Chebyshev's inequality,

$$\begin{aligned}\mathbf{P}\Big(M_{k_n} < -\varepsilon s_n(LLs_n^2)^{1/2}\Big) &\leqq \mathbf{P}\Big(S_{1n} < -\varepsilon s_n(LLs_n^2)^{1/2}\Big) \leqq \\ &\leqq (\varepsilon^2 LLs_n^2)^{-1} = o(1).\end{aligned} \tag{4.12}$$

Since $\varepsilon > 0$ is arbitrary in (4.11)–(4.12), these two relations validate (4.3) when $\beta_1 = 0$. □

COROLLARY 4.1. *Let $\mathcal{X}$ be t.i. and c.i.d. with*

$$\mathbf{E}X_{1n} = 0, \quad s_n^2 = \sum_{j=1}^{n}\mathbf{E}X_{1j}^2 \to \infty,$$

$k_n = (1+o(1))(Ls_n^2)^{\beta_1}(LLs_n^2)^{\beta_2}$ *where* $\beta_1 > 0$ *or* $\beta_1 = 0$, $\beta_2 \geqq 1$ *and either*

$$\inf_{n\geqq 1} \mathbf{E}X_{1n}^2 > 0, \quad \sup_{n\geqq 1} \mathbf{E}X_{1n}^2 (L|X_{1n}|)^{\beta_1}(LL|X|)^{\beta_2-1})L_m|x|)^{\eta} < \infty \tag{4.13}$$

for some $m \geqq 3$, $\eta > 0$ *or*

$$\sum_{j=1}^{n} \mathbf{E}X_{1j}^2 I\left[|X_{1j}| > \varepsilon s_j (LLs_j^2)^{-1/2}\right] = o\left(\frac{s_n^2 (LLs_n^2)^{1-\beta_2}}{(Ls_n^2)^{\beta_1}}\right), \quad \textit{all } \varepsilon > 0, \tag{4.14}$$

then (4.3) *and* (4.4) *hold.*

Proof. For k_n as specified, (4.1) holds and (4.14) coincides with (4.2). Thus, by Theorem 4.1, it is sufficient to verify that (4.13) implies (4.14). Now, under (4.13), $\mathbf{E}X_{1n}^2$ is bounded away from zero and infinity whence for all large j

$$\begin{aligned}\mathbf{E}X_{1j}^2 I\left[|X_{1j}| > \varepsilon s_j (LLs_j^2)^{-1/2}\right] &\leqq \frac{C}{(Ls_j^2)^{\beta_1}(L_2 s_j^2)^{\beta_2-1}(L_m s_j^2)^{\eta}} \\ &\leqq \frac{C'}{(Lj)^{\beta_1}(L_2 j)^{\beta_2-1}(L_m j)^{\eta}},\end{aligned}$$

implying

$$\sum_{j=1}^{n} \mathbf{E}X_{1j}^2 I\left[|X_{1j}| > \varepsilon s_j (LLs_j^2)^{-1/2}\right] = o\left(\frac{n(LLn)^{1-\beta_2}}{(Ln)^{\beta_1}}\right) = o\left(\frac{s_n^2(LLs_n^2)^{1-\beta_2}}{(Ls_n^2)^{\beta_1}}\right),$$

and establishing the corollary. □

Remark 4.1. If (3.1) and (3.3)′ hold, Theorem 4.1 and Corollary 4.1 are valid under (3.23)–(3.24) since, by the proof of Theorem 3.1, these conditions imply that $(s_n^2 LLs_n^2)^{-1/2} \max_{1\leqq i \leqq k_n} |T_{in}| \to 0$ a.s..

The next theorem gives conditions for equality to hold in (4.4).

Theorem 4.2. *Let* $\mathcal{X}$ *be t.i. and c.i.d. with*

$$\mathbf{E}X_{1n} = 0, \quad s_n^2 = \sum_{j=1}^{n} \mathbf{E}X_{1j}^2 \to \infty.$$

If (3.23), (3.24) *hold and* $(Lk_n)/(LLs_n^2) \to \beta_1 \in (0,\infty)$, *then*

$$\liminf_{n\to\infty} (2s_n^2 LLs_n^2)^{-1/2} M_{k_n} = \beta_1^{1/2} \quad \textit{a.s.}. \tag{4.15}$$

PROOF. Let S'_{1n}, s'_n, M'_{k_n} be as in the proof of Theorem 3.1 and (4.5) with

$$d_n = \bar{\varepsilon}_n s_n (LLs_n^2)^{-1/2}.$$

In view of Remark 4.1 and (4.6), it suffices to show that

$$\liminf_{n\to\infty} b_n^{-1} M'_{k_n} \geqq \beta_1^{1/2} \quad \text{a.s.,} \tag{4.16}$$

where $b_n = (2s_n^2 LLs_n^2)^{1/2}$. Define, for $\alpha \in (0,1)$, $n_1 = 1$ and

$$n_r = \sup\left\{n > n_{r-1} : s_n'^{\,2} \leqq (1-\alpha(r-1)^{-1+\alpha})^{-1} s'^2_{n_{r-1}}\right\} \text{ for } r \geqq 2. \tag{4.17}$$

Since for $r \geqq 3$,

$$s'^2_{n_r} \geqq s'^2_{n_{r-1}+1} > (1-\alpha(r-2)^{-1+\alpha})^{-1} s'^2_{n_{r-2}}$$

we have, for all large m,

$$s'^2_{n_{2m}} \geqq c\exp(\alpha' m^\alpha), \tag{4.18}$$

where $c > 0$ is an appropriate constant, and $\alpha' = \alpha/2^{1-\alpha}$. Moreover, by (4.17), for $r \geqq 2$,

$$g_r^2 := \left(1 - s'^2_{n_{r-1}}/s'^2_{n_r}\right)^{-1} \geqq \alpha(r-1)^{1-\alpha} \to \infty. \tag{4.19}$$

Set $k'_r = k_{n_{r-1}}$. For $n_{r-1} < n < n_r$, we have

$$M'_{k_n} \geqq M'_{k'_r} - \max_{1\leqq i\leqq k'_r} \max_{n_{r-1}<n\leqq n_r} \left|S'_{in} - S'_{in_{r-1}}\right|.$$

Hence, if it can be shown that

$$b_{n_r}^{-1} \max_{1\leqq i\leqq k'_r} \max_{n_{r-1}<n\leqq n_r} \left|S'_{in} - S'_{in_{r-1}}\right| \to 0 \tag{4.20}$$

and

$$\liminf_{n\to\infty} b_{n_r}^{-1} M'_{k'_r} \geqq \beta_1^{1/2} \quad \text{a.s.,} \tag{4.21}$$

it will follow that

$$\liminf_{n\to\infty} b_n^{-1} M'_{k_n} \geqq \liminf_{r\to\infty} b_{n_r}^{-1} M'_{k'_r} \geqq \beta_1^{1/2} \quad \text{a.s.,} \tag{4.22}$$

yielding (4.16). Now, recalling (4.19), we see that for any $\gamma \in (0,1)$ and all large r,

$$\begin{aligned}
&\mathbf{P}\Big(\max_{1\leqq i\leqq k_r'}\max_{n_{r-1}<n\leqq n_r}\Big|S_{in}'-S_{in_{r-1}}'\Big|>\gamma b_{n_r}\Big)\\
&\leqq k_r'\mathbf{P}\Big(\max_{n_{r-1}<n\leqq n_r}\Big|S_{1n}'-S_{1n_{r-1}}'\Big|>\gamma b_{n_r}\Big)\\
&\leqq k_r'\mathbf{P}\Big(\max_{n_{r-1}<n\leqq n_r}\Big|S_{1n}'-S_{1n_{r-1}}'\Big|>\\
&\qquad > \Big(s_{n_r}'^2-s_{n_{r-1}}'^2\Big)^{1/2}g_r\{2\gamma^2(1+o(1))LLs_{n_r}^2\}^{1/2}\Big)\\
&\leqq \exp\big\{(2\beta_1-\gamma^3g_r^2)LLs_{n_r}^2\big\}\leqq (Ls_{n_r}^2)^{-\gamma^4g_r^2}\leqq Cr^{-\alpha\gamma^4g_r^2},
\end{aligned}$$

by Lemma 10.2.1 of Chow and Teicher [2]. Since $g_r\to\infty$ as $r\to\infty$, this bound is summable for all $\gamma\in(0,1)$, whence the Borel–Cantelli lemma yields (4.20).

Furthermore, noting via (4.19) that $s_{n_r}'^2/s_n'^2\leqq s_{n_r}'^2/s_{n_{r-1}}'^2\leqq 1+o(1)$ for $n_{r-1}<n<n_r$, and making use of the observation that

$$LLs_{n_{r-1}}^2=(1+o(1))LLs_{n_r}^2,$$

we see that, for all large r,

$$\begin{aligned}
&\mathbf{P}\Big(M_{k_r'}'\leqq(1-\gamma)^2\beta_1^{1/2}b_{n_r}\Big)\\
&=\Big\{1-\mathbf{P}\Big(S_{1n}'>(1-\gamma)^2s_{n_r}'\{2\beta_1(1+o(1))LLs_{n_r}^2\}^{1/2}\Big)\Big\}^{[k_r']}\\
&\leqq\Big\{1-\mathbf{P}\Big(S_{1n}'>(1-\gamma)^2s_n'\{2\beta_1(1+o(1))LLs_{n_r}^2\}^{1/2}\Big)\Big\}^{[k_r']}\\
&\leqq\{1-C_\gamma\exp(-(1-\gamma)\beta_1LLs_{n_r}^2)\}^{[k_r']}\\
&\leqq\exp\{C_\gamma[k_r']\exp(-(1-\gamma)\beta_1LLs_{n_r}^2)\}\\
&\leqq\exp\{C_\gamma\exp((\beta_1\gamma+o(1))LLs_{n_r}^2)\}\\
&\leqq\exp\Big\{-C_\gamma(Ls_{n_r}^2)^{\gamma\beta_1/2}\Big\}\leqq\exp\Big\{-C_\gamma r^{\alpha\gamma\beta_1/2}\Big\},
\end{aligned}$$

which is summable for all $\gamma>0$, whence (4.21) follows by Borel–Cantelli and $\gamma\downarrow 0$. □

COROLLARY 4.2. *Let $\mathcal{X}$ be t.i. and c.i.d. with*

$$\mathbf{E}X_{1n}=0,\quad s_n^2=\sum_{j=1}^n\mathbf{E}X_{1j}^2\to\infty$$

and $s_{n+1}^2/s_n^2 = 1 + O(k_n^{-1})$. *If* (3.23), (3.24) *hold and*

$$(Lk_n)/(LLs_n^2) \downarrow \beta_1 \in (0,\infty),$$

then the almost sure limit set of $(2s_n^2 LLs_n^2)^{-1/2} M_{k_n}$ *is* $[\beta_1^{1/2}, (1+\beta_1)^{1/2}]$.

PROOF. By Theorems 3.2 and 4.2, the upper and lower a.s. limits of $(2s_n^2 LLs_n^2)^{-1/2} M_{k_n}$ are $(1+\beta_1)^{1/2}$ and $\beta_1^{1/2}$. Since (3.23) implies (2.18), the conclusion follows from Lemma 2.6 (i). □

COROLLARY 4.3. *Let $\mathcal{X}$ be t.i. and c.i.d. with*

$$\mathbf{E}X_{1n} = 0, \quad s_n^2 = \sum_{j=1}^{n} \mathbf{E}X_{1j}^2 \to \infty$$

and $\sigma_n^2 = \mathbf{E}X_{1n}^2 = o(s_n^2/LLs_n^2)$. *If* $(Lk_n)/L_3 s_n^2 \downarrow \beta_2 \in (0,\infty)$ *and* (2.3), (3.23), (3.24) *hold, then the almost sure limit set of* $(2s_n^2 LLs_n^2)^{-1/2} M_{k_n}$ *is* $[0,1]$.

PROOF. Corollary 3.1 guarantees an upper limit of $(2s_n^2 LLs_n^2)^{-1/2} M_{k_n}$ a.s. equal to 1, while Theorem 4.1 in conjunction with Remark 4.1 ensures non-positivity of the corresponding lower limit. In view of (3.24) and (2.3), Corollary 2.1 (ii) asserts the non-negativity of the lower limit. Thus, the lower limit is a.s. equal to 0, whence Lemma 2.6 yields the conclusion, recalling that (3.23) implies (2.15). □

In the case of large k_n, the next theorem upgrades the behaviour of the normalized M_{k_n} to almost sure convergence.

THEOREM 4.3. *Let $\mathcal{X}$ be t.i. and c.i.d. with*

$$\mathbf{E}X_{1n} = 0, \quad s_n^2 = \sum_{j=1}^{n} \mathbf{E}X_{1j}^2 \to \infty,$$

and let k_n satisfy $s_n^2/Lk_n \uparrow \infty$ *and* $(Lk_n)/LLs_n^2 \to \infty$. *If, for all* $\varepsilon > 0$,

$$\sum_{n=1}^{\infty} k_n \mathbf{P}\left(|X_{1n}| > \varepsilon s_n (Lk_n)^{-1/2}\right) < \infty, \tag{4.23}$$

and

$$\sum_{j=1}^{n} \mathbf{E}X_{1j}^2 I\left[|X_{1j}| > \varepsilon s_j (Lk_j)^{-1/2}\right] = o(s_n^2), \tag{4.24}$$

then

$$\lim_{n\to\infty} (2s_n^2 Lk_n)^{-1/2} M_{k_n} = 1 \quad \textit{in probability.} \tag{4.25}$$

Moreover, if $s_n^2 \geqq c_0 n^\alpha$ *for some* $\alpha > 0$ *and* $c_0 > 0$*, then*

$$\lim_{n\to\infty} (2s_n^2 Lk_n)^{-1/2} M_{k_n} = 1 \quad a.s.. \tag{4.26}$$

Proof. Let S'_{in}, T_{in}, M'_{k_n} and s'_n be as in the proofs of Theorems 3.1 and 4.1 (see e.g. (3.10)–(3.13) and (4.5)), where now $d_n = \varepsilon_n s_n (Lk_n)^{-1/2}$. According to (3.11)–(3.12) in the proof of Theorem 3.1, if $b_n = (2s_n^2 Lk_n)^{1/2}$ then

$$b_n^{-1} \max_{1\leqq i\leqq k_n} |T_{in}| \to 0 \text{ a.s.}. \tag{4.27}$$

Hence, in view of (4.6), to establish (4.25), it suffices to verify that

$$b_n^{-1} M'_{k_n} \to 1 \text{ in probability.}$$

Now (4.24) guarantees $s_n = (1+o(1))s'_n$ and so, as in the proof of Theorem 3.1, we have

$$|X'_{ij}| \leqq 2d_n \leqq 3\overline{\varepsilon}_n s'_n (Lk_n)^{-1/2} \leqq \varepsilon_n s'_n (Lk_n)^{-1/2}, \; 1 \leqq j \leqq k_n,$$

for all large n, whence in analogy with (3.16), for $\lambda > \rho > 1$,

$$k_n \mathbf{P}\Big(S'_{1n} > \lambda^{1/2} b_n\Big) = k_n \mathbf{P}\Big(S'_{1n} > (1+o(1))s'_n (2\lambda Lk_n)^{1/2}\Big)$$
$$\leqq k_n \exp\Big(-\varepsilon_n^{-2} h((2\rho)^{1/2}\varepsilon_n) Lk_n\Big) \leqq k_n^{1-\rho} = o(1).$$

This, in turn, implies that, for any fixed $\lambda > 1$,

$$\mathbf{P}\Big(M'_{k_n} > \lambda^{1/2} b_n\Big) = o(1). \tag{4.29}$$

Likewise, and recalling the argument preceding (3.21), we see that, for any fixed $\gamma \in (0,1)$, there exists a constant $C_\gamma > 0$ such that, as $n \to \infty$,

$$\begin{aligned} &\mathbf{P}(M'_{k_n} \leqq (1-\gamma)^2 b_n) \\ &= \Big\{1 - \mathbf{P}\Big(S'_{1n} > (1-\gamma)^2 s'_n \{2(1+o(1))Lk_n\}^{1/2}\Big)\Big\}^{[k_n]} \\ &\leqq \exp\big(-[k_n] C_\gamma \exp\{-(1-\gamma)^2 Lk_n\}\big) \leqq \exp(-C_\gamma k_n^\gamma) = o(1). \end{aligned} \tag{4.30}$$

Thus, by combining (4.29) and (4.30), we obtain (4.25) by letting $\lambda \downarrow 1$ and $\gamma \downarrow 0$.

Moreover, since $(Lk_n)/LLs_n^2 \to \infty$ is equivalent to $r_n := k_n/(Ls_n^2)^\beta \to \infty$ for all $\beta > 0$, under the assumption that $s_n^2 \geqq c_0 n^\alpha$, we see via (4.30) that for all large n

$$\mathbf{P}(M'_{k_n} \leqq (1-\gamma)^2 b_n) \leqq \exp\Big(-C_\gamma r_n^\gamma (Ls_n^2)^{\beta\gamma}\Big) \leqq \exp(-(Ls_n^2)^{\beta\gamma}) \leqq n^{-2},$$

by choosing $\beta = 2/\gamma$. Thus the Borel–Cantelli lemma ensures, as $\gamma \downarrow 0$, $\liminf_{n\to\infty} b_n^{-1} M'_{k_n} \geqq 1$ a.s., which, in turn, guarantees $\liminf_{n\to\infty} b_n^{-1} M_{k_n} \geqq 1$ a.s.. Since (3.8) in Theorem 3.1 asserts that $\limsup_{n\to\infty} b_n^{-1} M_{k_n} \leqq 1$ a.s., the combination of these two relation yields (4.26). □

5. The totally i.i.d. case

As is to be expected, the prior results simplify when $\mathcal{X}$ is t.i.i.d. Let for $\eta > 0$ and $m \geqq 2$

$$g_{m,\eta}(x) = L_{m+1}^{1+\eta}|x| \prod_{i=2}^{m} L_i|x|. \tag{5.1}$$

THEOREM 5.1. *Let $\mathcal{X}$ be t.i.i.d. with $\mathbf{E}X = 0$, $\mathbf{E}X^2 = 1$. If*
(i) $(Lk_n)/(LLn) \to \beta_1 \in (0, \infty)$ *and*

$$\sum_{n=1}^{\infty} k_n \mathbf{P}\left(|X| > \varepsilon \left(\frac{n}{LLn}\right)^{1/2}\right) < \infty, \qquad \textit{all } \varepsilon > 0, \tag{5.2}$$

or
(ii) $k_n = (Ln)^{\beta_1}(LLn)^{\beta_2}(1 + o(1))$, $\beta_1 \geqq 0$, $\beta_1 \vee \beta_2 > 0$ *and*

$$\mathbf{E}X^2(L|X|)^{1+\beta_1}(LL|X|)^{1+\beta_2} g_{m,\eta}(X) < \infty \textit{ for some } m \geqq 2 \textit{ and } \eta \geqq 0, \tag{5.3}$$

then

$$\limsup_{n\to\infty} (2nLLn)^{-1/2} M_{k_n} = (1 + \beta_1)^{1/2} \quad a.s.. \tag{5.4}$$

PROOF. Since (3.24) holds in the t.i.i.d. case, (i) follows from Theorem 3.2. Under (5.3),

$$\mathbf{P}\left(|X| > \varepsilon \left(\frac{n}{LLn}\right)^{1/2}\right) \leqq \frac{CLLn}{n(Ln)^{1+\beta_1}(LLn)^{1+\beta_2}(L_{m+1}n)^{1+\eta} \prod_{i=2}^{m} L_i n},$$

whence (5.2) holds and Case (ii) follows via Theorem 3.2 and Corollary 3.1. □

REMARK 5.1. When $k_n = \overline{k}_n(\beta_1, \beta_2)$ is as in (3.28), Corollary 3.2 reveals that (5.3) can be weakened to $\mathcal{H}(1 + \beta_1, 2 + \beta_2) < \infty$. Also in Case (ii) of Theorem 5.1 when $k_n = (1 + o(1))(LLn)^{\beta_2}$ with $\beta_2 > 0$, (5.4) holds with $\beta_1 = 0$.

THEOREM 5.2. *Let $\mathcal{X}$ be t.i.i.d. with $\mathbf{E}X = 0$, $\mathbf{E}X^2 = 1$. If*
(i) $(Lk_n)/(LLn) \to \beta_1 \in [0, \infty)$ *and*

$$\sum_{j=1}^{n} \mathbf{E}X^2 I\left[|X| > \varepsilon \left(\frac{j}{LLj}\right)^{1/2}\right] = o\left(\frac{nLLn}{k_n}\right), \tag{5.5}$$

or
(ii) $k_n = (1 + o(1))(Ln)^{\beta_1}(LLn)^{\beta_2}$, $\beta_1 \geqq 0$, $\beta_1 \vee \beta_2 > 0$ *and*

$$\mathbf{E}X^2(L|X|)^{\beta_1}(LL|X|)^{\beta_2 - 1}(L_{m+1}|X|)^{\eta} < \infty \tag{5.6}$$

for some $m \geqq 2$, *and* $\eta > 0$, *then*

$$\lim_{n\to\infty} (2nLLn)^{-1/2} M_{k_n} = \beta_1^{1/2} \quad \textit{in probability,} \tag{5.7}$$

and

$$\liminf_{n\to\infty} (2nLLn)^{-1/2} M_{k_n} \leqq \beta_1^{1/2} \quad \textit{a.s.}. \tag{5.8}$$

Moreover, when $\beta_1 = 0$, *equality holds in* (5.8).

PROOF. This follows immediately from Theorem 4.1 and Corollary 4.1. The final remark emanates from Corollary 2.1 (i). □

THEOREM 5.3. *Let* $\mathcal{X}$ *be t.i.i.d. with* $\mathbf{E}X = 0$, $\mathbf{E}X^2 = 1$. *If*
(i) $(Lk_n)/(LLn) \downarrow \beta_1 \in (0, \infty)$ *and*

$$\sum_{j=1}^{n} k_n \mathbf{P}\left(|X| > \varepsilon \left(\frac{n}{LLn} \right)^{1/2} \right) < \infty \quad \textit{for all } \varepsilon > 0, \tag{5.9}$$

or
(ii) $k_n = (1 + o(1))(Ln)^{\beta_1}(LLn)^{\beta_2}$, $\beta_1 \geqq 0$, $\beta_1 \vee \beta_2 > 0$ *and*

$$\mathbf{E}X^2 (L|X|)^{\beta_1+1} (LL|X|)^{\beta_2+1} g_{m,\eta}(X) < \infty \quad \textit{for some } m \geqq 2,\ \eta > 0, \tag{5.10}$$

then the almost sure limit set of $(2nLLn)^{-1/2} M_{k_n}$ *is* $\left[\beta_1^{1/2}, (1+\beta_1)^{1/2}\right]$.

PROOF. According to Theorems 4.2 and 5.1, the upper and lower limits are $(1 + \beta_1)^{1/2}$ and $\beta_1^{1/2}$ a.s.. Thus, under (i), the conclusion follows from Lemma 2.6 (i). In Case (ii), Remark 5.1 and Theorem 5.2 (ii) ensure the upper and lower limits even when $\beta_1 = 0$, and since (5.10) implies $\mathcal{H}(\beta_1, \beta_2 - 1) < \infty$, Lemma 2.6 (i) guarantees the result for $\beta_1 \geqq 0$. □

THEOREM 5.4. *Let* $\mathcal{X}$ *be t.i.i.d. with* $\mathbf{E}X = 0$, $\mathbf{E}X^2 = 1$ *and let* k_n *satisfy* $n/(Lk_n) \uparrow \infty$ *and* $(Lk_n)/LLn \to \infty$. *If for all* $\varepsilon > 0$,

$$\sum_{j=1}^{n} k_n \mathbf{P}\left(|X| > \varepsilon \left(\frac{n}{Lk_n} \right)^{1/2} \right) < \infty, \tag{5.11}$$

then

$$\lim_{n\to\infty} (2nLk_n)^{-1/2} M_{k_n} = 1 \quad \textit{a.s.}. \tag{5.12}$$

PROOF. Since $s_n^2 = n$, (5.11) is equivalent to (4.23) while (4.23) is automatic. Therefore, the result is an immediate consequence of Theorem 4.3.
□

COROLLARY 5.1. *Let $\mathcal{X}$ be t.i.i.d. with* $\mathbf{E}X = 0$, $\mathbf{E}X^2 = 1$ *and*

$$\mathbf{E}|X|^{2\alpha+2}(L|X|)^{\alpha+2}g_{m,\eta}(X) < \infty, \tag{5.13}$$

for some $\alpha > 0$, $m \geqq 2$ *and* $\eta > 0$. *Then*

$$\lim_{n\to\infty} (2\alpha nLn)^{-1/2} M_{n^\alpha} = 1 \quad a.s.. \tag{5.14}$$

PROOF. When $k_n = n^\alpha$, $\alpha > 0$, (5.11) holds. Therefore, the result follows from Theorem 5.4. □

REMARK 5.2. For $\alpha = 1$, Corollary 5.1 and (1.3) imply that, if

$$\mathbf{E}X^4(L|X|)^{3+\varepsilon} < \infty \text{ for some } \varepsilon > 0,$$

then

$$\limsup_{n\to\infty} (2nLn)^{-1/2} S'_{nn} = \lim_{n\to\infty} (2nLn)^{-1/2} M_n = 1 \quad \text{a.s.}, \tag{5.15}$$

which shows that (5.14) is in agreement with the characterizations of the upper class of S_{nn} due to Cramér [3] and Esseen [11] (see e.g. (1.3)). When $\alpha = 1$, (5.13) is slightly stronger than the Cramér–Esseen condition $\mathbf{E}X^4 < \infty$ (which according to Esseen [11] cannot be weakened), but results in a stronger conclusion.

6. LIL's for arrays via invariance principles

In this section, we assume that $\mathcal{X}$ is t.i.i.d. and X is a random variable following the distribution of X_{ij} for $i \geqq 1$ and $j \geqq 1$. We assume further that

$$\mathbf{E}X = 0 \quad \text{and} \quad \mathbf{E}X^2 = 1. \tag{6.1}$$

The following Fact 3 follows from Theorem 6.5 of Major [22] (see also Komlós, Major and Tusnády [18], [19] and Theorem 2.6.7, p. 110 in Csörgő and Révész [4]), taken with $H(x) = x^{2+\delta}$, and applied to each of the sequences $\{S_{ij},\ j \geqq 1\}$ for $i = 1, 2, \ldots$.

FACT 3. *Assume that* (6.1) *holds, and that* $\mathbf{E}(|X|^{2+\delta}) < \infty$ *for some* $\delta > 0$. *Then, there exist two constants* $c_1 > 0$ *and* $c_2 > 0$ *depending upon the distribution of* X *only, such that the following property holds. We may define* $\mathcal{X}$ *on* $(\Omega, \mathcal{A}, \mathbf{P})$ *jointly with a sequence of independent standard Wiener processes* $W_1(\cdot), W_2(\cdot), \ldots$, *such that, provided* $n^{1/(2+\delta)} < x_n < c_1(nLn)^{1/2}$,

$$\mathbf{P}\left(\max_{1\leqq j\leqq n} |S_{ij} - W_i(j)| \geqq x_n\right) \leqq c_2 n x_n^{-2-\delta} \quad \textit{for } i \geqq 1 \textit{ and } n \geqq 1. \tag{6.2}$$

Let $\{W_n(t),\ t \geqq 0\}$, $n = 1, 2, \ldots$ be as in Fact 3, and set $\mu_{k_n} = \mu_{k_n}(n) = \max_{1\leqq i\leqq k_n} W_i(n)$. The next lemma shows that the existence of higher order moments implies that $\mu_{k_n} = M_{k_n} + o(n^{1/2})$ a.s..

LEMMA 6.1. *If* $\mathbf{E}|X|^{2+\delta} < \infty$ *and* $\limsup_{n\to\infty}(Lk_n)/Ln < \delta$ *for some* $\delta > 0$, *then*

$$\lim_{n\to\infty} n^{-1/2}|M_{k_n} - \mu_{k_n}| = 0 \quad a.s.. \tag{6.3}$$

PROOF. Let $n_m = 2^m$ for $m \geqq 0$. For all large n, we have $k_n \leqq n^\gamma$, where γ is a constant such that $0 < \gamma < \delta$. It follows from (6.2) that, for all m sufficiently large and $\varepsilon > 0$,

$$\mathbf{P}\left(\bigcup_{n_{m-1}<n\leqq n_m} \left\{|M_{k_n} - \mu_{k_n}| > \varepsilon n^{1/2}\right\}\right)$$

$$\leqq \mathbf{P}\left(\bigcup_{n_{m-1}<n\leqq n_m} \left\{\max_{1\leqq j\leqq n_m} |S_{ij} - W_i(j)| > \varepsilon n_m^{1/2}\right\}\right) \leqq k_{n_m} c_2 n_m \{\varepsilon n_m^{1/2}\}^{-2-\delta}$$

$$\leqq \exp\left\{-m\left(\delta - \frac{1}{2}(\delta-\gamma)\right)\log 2 + Lk_{n_m}\right\} \leqq \exp\left\{-\frac{1}{2}m(\delta-\gamma)\log 2\right\},$$

which is summable in m. By the Borel–Cantelli lemma and letting $\varepsilon \downarrow 0$, we obtain (6.3). □

We next consider for each $n \geqq 1$ the set of random functions defined by

$$\mathcal{F}_n = \left\{(2nLLn)^{-1/2} W_i(n\mathbf{I}) : 1 \leqq i \leqq k_n\right\}, \tag{6.4}$$

where $\mathbf{I}$ is the identity mapping of $[0,1]$ onto itself. Let $\mathcal{B}_0(0,1)$ denote the set of all bounded functions f on $[0,1]$ with $f(0) = 0$, endowed with the topology defined by the *sup-norm* $\|f\| = \sup_{0\leqq t\leqq 1} |f(T)|$. For any $f \in \mathcal{B}_0(0,1)$ and $\varepsilon > 0$, set $N_\varepsilon(f) = \{g \in \mathcal{B}(0,1) : \|f-g\| < \varepsilon\}$, and, for any $B \subseteqq B_0(0,1)$, set $B^\varepsilon = \bigcup_{f\in B} N_\varepsilon(f)$. For each $c \geqq 0$, let $\mathcal{S}_c = \{f \in \mathcal{B}_0(0,1) : |f|_{\mathcal{H}} \leqq c\}$, where

$$|f|_{\mathcal{H}} = \left\{\int_0^1 \dot{f}(t)^2 dt\right\}^{1/2} \quad \text{if } f \text{ is absolutely continuous on } [0,1] \text{ with } \dot{f} = df/dx,$$

$$|f|_{\mathcal{H}} = \infty \quad \text{otherwise.} \tag{6.5}$$

Our arguments will rely in part on the following fact, due to Deheuvels and Révész [8].

FACT 4. *Let $\varepsilon > 0$ be arbitrary, and assume that $(Lk_n)/LLn \to \beta_1 \in \in [0,\infty)$. Then*

$$\lim_{n\to\infty} \mathbf{P}(\mathcal{S}_{\beta_1} \subseteqq \mathcal{F}_n^\varepsilon \text{ and } \mathcal{F}_n \subseteqq \mathcal{S}_{\beta_1}^\varepsilon) = 1, \tag{6.6}$$

and there exists almost surely an $n'_\varepsilon < \infty$ such that, for all $n \geqq n'_\varepsilon$,

$$\mathcal{S}_{\beta_1} \subseteqq \mathcal{F}_n^\varepsilon \quad \text{and} \quad \mathcal{F}_n \subset \mathcal{S}_{1+\beta_1}^\varepsilon. \tag{6.7}$$

Moreover, for any $f \in \mathcal{S}_1$ with $\beta_1 < d := |f|_{\mathcal{H}}^2 < 1 + \beta_1$,

$$\mathbf{P}\big(f \in \mathcal{F}_n^\varepsilon \subset \mathcal{S}_d^{2\varepsilon} \text{ i.o.}\big) = 1. \tag{6.8}$$

PROOF. See e.g. Theorem 1.1 and Remark 2.3 in Deheuvels and Révész [8]. □

An easy consequence of Fact 4 is stated in the following proposition.

PROPOSITION 6.1. *Assume that $(Lk_n)/LLn \to \beta_1 \in [0,\infty)$ and $k_n \to \infty$. Then, the almost sure limit set of $\{(2nLLn)^{-1/2}k_n\}$ is equal to $[\beta_1^{1/2}, (1+\beta_1)^{1/2}]$. Moreover,*

$$(2nLLn)^{-1/2}\mu_{k_n} \to \beta_1^{1/2} \quad \text{in probability.} \tag{6.9}$$

PROOF. Since $(2nLLn)^{-1/2}\mu_{k_n} = \sup_{f\in\mathcal{F}_n} f(1)$, by (6.8) and the triangle inequality, we have almost surely for any fixed $\varepsilon > 0$ and all large n,

$$-\varepsilon + \sup_{f\in\mathcal{S}_{\beta_1}} f(1) \leqq \sup_{f\in\mathcal{F}_n} f(1) \leqq \varepsilon + \sup_{f\in\mathcal{S}_{1+\beta_1}} f(1). \tag{6.10}$$

Since $\sup_{f\in\mathcal{S}_{\beta_1}} f(1) = d^{1/2}$, by letting $\varepsilon \downarrow 0$ in (6.10), we obtain that the almost sure limit set $\mathcal{L}$ of $\{(2nLLn)^{-1/2}\mu_{k_n}\}$ is included into $\left[\beta_1^{1/2}, (1+\beta_1)^{1/2}\right]$. To show that $\mathcal{L} = \left[\beta_1^{1/2}, (1+\beta_1)^{1/2}\right]$, we choose an arbitrary $d \in \left(\beta_1^{1/2}, (1+\beta_1)^{1/2}\right)$ and set $f = d^{1/2}\mathbf{I} \in \mathcal{S}_d$. By (6.8), for any $\varepsilon > 0$, there exists almost surely an unbounded sequence along which $f \in \mathcal{F}_n^\varepsilon$ and $\mathcal{F}_n \subseteqq \mathcal{S}_d^\varepsilon$, and hence

$$-\varepsilon + f(1) = -\varepsilon + d \leqq \sup_{f\in\mathcal{F}_n} f(1) \leqq \varepsilon + \sup_{f\in\mathcal{S}_d} f(1) = \varepsilon + d. \tag{6.11}$$

Letting $\varepsilon \downarrow 0$ in (6.11) shows that $d \in \mathcal{L}$. Since $\mathcal{L} \subseteqq \mathcal{B}_0(0,1)$ is compact, a simple argument shows that $\mathcal{L} = \left[\beta_1^{1/2}, (1+\beta_1)^{1/2}\right]$. The proof of (6.9) follows likewise from (6.6). □

PROPOSITION 6.2. *Let $\mathcal{X}$ be t.i.i.d., let* (1.1) *hold and assume that there exists a* $\delta > 0$ *such that* $\mathbf{E}|X|^{2+\delta} < \infty$. *If*

$$(6.12) \qquad (Lk_n)/LLn \to \beta_1 \in [0, \infty) \text{ and } k_n \to \infty,$$

then the almost sure limit set of $\{(2nLLn)^{-1/2}M_{k_n}\}$ *is* $\left[\beta_1^{1/2}, (1+\beta_1)^{1/2}\right]$ *and*

$$(6.13) \qquad (2nLLn)^{-1/2}M_{k_n} \to \beta_1^{1/2} \text{ in probability.}$$

PROOF. Combine Lemma 6.1 and Proposition 6.1. □

The just-given simple proof of Proposition 6.2 sheds light on the mechanism which generates the limit set of $(2nLLn)^{-1/2}M_{k_n}$, but brings only minor improvements to the results of Section 5 and necessitates the unnecessarily restrictive assumption that $\mathbf{E}|X|^{2+\delta} < \infty$ for some $\delta > 0$. When one assumes only that $H(\beta_1, \beta_2) = \mathbf{E}X^2(L|X|)^{\beta_1}(LL|X|)^{\beta_2} < \infty$, a different argument is needed. The following result follows this line of thought and improves upon Theorem 5.3.

THEOREM 6.1. *If $\mathcal{X}$ is t.i.i.d. with* $\mathbf{E}X = 0$, $\mathbf{E}X^2 = 1$, $k_n = (Ln)^{\beta_1}(LLn)^{\beta_2}$ *and*

$$(6.14) \qquad H(\beta_1, \beta_2) < \infty, \text{ for } \beta_1 \geqq 0 \text{ and } \beta_1 \vee \beta_2 > 0,$$

then the almost sure limit set of $(2nLLn)^{-1/2}M_{k_n}$ *is* $\left[\beta_1^{1/2}, (1+\beta_1)^{1/2}\right]$.

In the remainder of this section, we prove Theorem 6.1. The following fact gives an essential clue to the solution of this problem.

FACT 5. *Under* (6.1), *we may define $\mathcal{X}$ on* $(\Omega, \mathcal{A}, \mathbf{P})$ *jointly with a sequence of independent standard Wiener processes* $W_1(\cdot), W_2(\cdot), \dots,$ *and a t.i.i.d. array* $\mathcal{Z} = \{Z_{ij},\ i \geqq 1, j \geqq 1\}$ *such that the following properties hold. If* $Z = Z_{11}$, *and* $V_{in} = \sum_{j=1}^{n} Z_{ij}$ *for* $i \geqq 1$ *and* $n \geqq 1$, *then*

$$(6.15) \qquad S_{in} = W_i(V_{in}) \text{ for } i \geqq 1 \text{ and } n \geqq 1,$$

$$(6.16) \qquad EZ = 1 \text{ and } \mathbf{P}(Z < 0) = 0.$$

PROOF. Apply the Skorokhod [24] embedding scheme to $\{S_{in}, n \geqq 1\}$ for $i = 1, 2, \dots$. □

We assume from now on that Z, $\mathcal{Z} = \{Z_{ij},\ i \geqq 1, j \geqq 1\}$ and $W_1, W_2, \dots$ are as in Fact 5. The following lemma will play an instrumental role in the sequel.

LEMMA 6.2. *Assume that* (6.1), (6.6) *and* (6.16) *hold and that*

$$\lim_{n\to\infty}\left\{\max_{1\leqq i\leqq k_n}|n^{-1}V_{in}-1|\right\}=0 \quad a.s.. \tag{6.17}$$

Let $G_n=\{(2nLLn)^{-1/2}W_i(V_{in}\mathbf{I}) : 1\leqq i\leqq n\}$. *Then, we have a.s. for any* $\varepsilon>0$ *and all large* n

$$\mathcal{F}_n\subseteq\mathcal{G}_n^{\varepsilon} \quad and \quad \mathcal{G}_n\subseteq\mathcal{F}_n^{2\varepsilon}. \tag{6.18}$$

PROOF. Recall (6.4), and let $\theta>0$ be a constant which will be specified later on. By (6.7), $\mathcal{F}_{2n}\subseteq\mathcal{S}^{\theta}_{1+\beta_1}$ for all $n\geqq\frac{1}{2}n'_{\theta}$, whence, by the triangle inequality, for any $0\leqq\rho<2$ and $f\in\mathcal{F}_{2n}$,

$$\begin{aligned}&\sup_{0\leqq t\leqq 1}\left|f\left(\rho\frac{t}{2}\right)-f\left(\frac{t}{2}\right)\right|\\ &\leqq 2\theta+\sup_{0\leqq t\leqq 1}\sup_{g\in\mathcal{S}_1}\left|g\left(\rho\frac{t}{2}\right)-g\left(\frac{t}{2}\right)\right|\leqq 2\theta+2^{-1/2}|\rho-1|^{1/2}.\end{aligned} \tag{6.19}$$

The Schwarz inequality and (6.5) are used here to show that, for all $0\leqq u\leqq$ $\leqq v\leqq 1$ and $g\in\mathcal{S}_1$,

$$|g(v)-g(u)|=\left|\int_u^v\dot g(t)dt\right|\leqq|v-u|^{1/2}\int_u^v\dot g^2(t)dt\leqq|v-u|^{1/2}. \tag{6.20}$$

Fix $\varepsilon\in(0,1)$ and $\theta=\varepsilon/8$ in (6.19). By (6.17) there exists a.s. an $m'_{\varepsilon}<\infty$ such that for all $n\geqq m'_{\varepsilon}$

$$\max_{1\leqq i\leqq k_{2n}}|n^{-1}V_{in}-1|\leqq\varepsilon^2/16. \tag{6.21}$$

Since $f=(4nLL(2n))^{-1/2}W_i(2n\mathbf{I})\in\mathcal{F}_{2n}$ for $1\leqq i\leqq k_n$, by (6.19) taken with $\rho=n^{-1}V_{in}$, and making use of (6.21) which implies $0\leqq\rho\leqq 2$, we see that, for all $n\geqq\max\left(\frac{1}{2}n'_{\theta},m'_{\varepsilon}\right)$,

$$\begin{aligned}&(4nLL(2n))^{-1/2}\max_{1\leqq i\leqq k_{2n}}\sup_{0\leqq t\leqq 1}|W_i(nt)-W_i(V_{in}t)|\\ &\leqq 2\left(\frac{\varepsilon}{8}\right)+2^{-1/2}\left(\frac{\varepsilon^2}{16}\right)^{1/2}<\frac{\varepsilon}{2}.\end{aligned}$$

This, in turn, implies that, almost surely for all large n,

$$(2nLLn)^{-1/2}\max_{1\leqq i\leqq k_n}\sup_{0\leqq t\leqq 1}|W_i(nt)-W_i(V_{in}t)|\leqq\frac{\varepsilon}{2}\left(\frac{2LL(2n)}{LLn}\right)^{1/2}\leqq\varepsilon,$$

whence (6.18). □

Facts 4–5, Proposition 6.1 and Lemma 6.2 show that the conclusion of Proposition 6.2 holds if $\mathcal{Z}$ satisfies (6.17). The next proposition establishes sufficient conditions for (6.17) and has interest in itself.

PROPOSITION 6.3. *Assume that* $k_n = (Ln)^{\beta_1}(LLn)^{\beta_2}$, $\beta_1 \geqq 0$, $\beta_1 \vee \beta_2 > 0$ *and that*

$$\mathbf{E}Z(LZ)^{\beta_1}(LLZ)^{\beta_2} < \infty. \tag{6.22}$$

Then,

$$\lim_{n\to\infty}\left\{\max_{1\leqq i\leqq k_n} |n^{-1}V_{in} - 1|\right\} = 0 \quad a.s.. \tag{6.23}$$

PROOF. The assumption (6.22) implies that

$$\begin{aligned}
\sum_{n=1}^{\infty}\sum_{i=1}^{k_n}\mathbf{P}(Z_{in}\geqq n) &= \sum_{i=1}^{\infty}k_n\mathbf{P}(Z\geqq n) = \sum_{n=1}^{\infty}k_n\sum_{j=n}^{\infty}\mathbf{P}(j\leqq Z<j+1)\\
&= \sum_{j=1}^{\infty}\mathbf{P}(j\leqq Z<j+1)\sum_{n=1}^{j}k_n = \sum_{j=1}^{\infty}j(Lj)^{\beta_1}(LLj)^{\beta_2}\mathbf{P}(j\leqq Z<j+1)\\
&= O(\mathbf{E}Z(LZ)^{\beta_1}(LLZ)^{\beta_2}) < \infty.
\end{aligned} \tag{6.24}$$

Let

$$Z_{in}^* = Z_{in}I\left[|Z_{in}|\leqq n\right], \text{ and } V_{in}^* = \sum_{j=1}^{n}Z_{ij}^*.$$

By (6.24), the Borel–Cantelli lemma implies that $\max_{1\leqq i\leqq k_n} n^{-1}|V_{in} - V_{in}^*| \to 0$ a.s.. By (6.16),

$$\mathbf{E}Z_{in}^* = \mathbf{E}ZI\left[|Z|\leqq n\right] \to \mathbf{E}Z = 1, \quad \text{and hence,} \quad \max_{1\leqq i\leqq k_n} |n^{-1}\mathbf{E}V_{in}^* - 1| \to 0.$$

Therefore, all we need for (6.23) is to show that

$$\max_{1\leqq i\leqq k_n} n^{-1}|V_{in} - \mathbf{E}V_{in}^*| \to 0 \ \text{ a.s.}. \tag{6.25}$$

Let

$$\Delta_{mi} = \max_{2^{m-1}<n\leqq 2^m} n^{-1}|V_{in} - \mathbf{E}V_{in}^*|$$

and

$$\Gamma_{mi} = \max_{n>2^{m-1}} \Delta_{ni}.$$

Set

$$K(z) = (LZ)^{\beta_1}(LLZ)^{\beta_2},$$

and note for further use that $k_n = K(n)$. By Kolmogorov's inequality,

$$\begin{aligned}
&\mathbf{P}\left(\max_{1\leqq i\leqq k_{2N}} \Gamma_{ni} \geqq \varepsilon\right) \leqq \sum_{m=N}^{\infty} \mathbf{P}\left(\max_{1\leqq i\leqq k_{2N}} \Delta_{ni} \geqq \varepsilon\right)\\
&\leqq \sum_{m=N}^{\infty} k_{2N}\mathbf{P}(\Delta_{m1} \leqq \varepsilon)\\
&\leqq \sum_{m=N}^{\infty} k_{2N}\mathbf{P}\left(\max_{1\leqq n\leqq 2^m} |V_{1n}^* - \mathbf{E}V_{1n}^*| \geqq \varepsilon 2^{m-1}\right)\\
&\leqq \sum_{m=N}^{\infty} k_{2N}(\varepsilon 2^{m-1})^{-2} \sum_{1\leqq n\leqq 2^m} \mathbf{E}(Z_{1n}^* - \mathbf{E}Z_{1n}^*)^2\\
(6.26)\qquad &\leqq 4\varepsilon^{-2} \sum_{m=N}^{\infty} k_{2N}2^{-2m} \sum_{1\leqq n\leqq 2^m} \mathbf{E}Z^2 I[Z\leqq n]\\
&\leqq \frac{16}{3}\varepsilon^{-2}\left\{k_{2N}2^{-2N}\sum_{n=1}^{2^N} \mathbf{E}Z^2 I[Z\leqq n]\right.\\
&\qquad\left. +4\sum_{n=2^N+1}^{\infty} n^{-2}k_n \mathbf{E}Z^2 I[Z\leqq n]\right\}\\
&=: \frac{16}{3}\varepsilon^{-2}\{A_n + 4B_n\}.
\end{aligned}$$

Observe that B_N is the remainder term of the series

$$\begin{aligned}
&\sum_{n=1}^{\infty} n^{-2}k_n \mathbf{E}Z^2 I[Z\leqq n] \leqq \sum_{n=1}^{\infty} n^{-2}k_n \sum_{j=1}^{n} \mathbf{E}Z^2 I[j-1<Z\leqq j]\\
(6.27)\qquad &\leqq C\sum_{j=1}^{\infty} \mathbf{E}ZK(Z)I[j-1<Z\leqq j](j/K(j))\sum_{n=j}^{\infty} n^{-2}K(n)\\
&= O\left(Z(LZ)^{\beta_1}(LLZ)^{\beta_2}\right) < \infty,
\end{aligned}$$

and therefore $B_N = o(1)$. The Kronecker lemma asserts that if $\sum_{n=1}^{\infty} a_n <$ $<\infty$, and if $q_n \uparrow \infty$, with $a_n \geqq 0$, $q_n \geqq 0$ for $n \geqq 1$, then, $q_n^{-1}\sum_{j=1}^{n} a_j q_j \to 0$. By (6.27), we may apply this lemma with $a_n = n^{-2}k_n \mathbf{E}Z^2 I[Z\leqq n]$ and $q_n =$ $= n^2 k_n^{-1}$ to obtain that $A_N = k_{2N}2^{-2N}\sum_{n=1}^{2^N} \mathbf{E}Z^2 I[Z\leqq n] = o(1)$. Thus, by (6.26), $\mathbf{P}\left(\max_{1\leqq i\leqq k_{2N}} \Gamma_{Ni} \geqq \varepsilon\right) \to 0$, which, by letting $\varepsilon \downarrow 0$, suffices for (6.23).

□

The problem which remains to be solved is to establish conditions in terms of X which ensure (6.22). It is known (see e.g. Lemma A.2, p. 272 of Hall and Heyde [16]) that, for each $r \geqq 1$,

$$\mathbf{E}Z^r \leqq c_r \mathbf{E}|X|^{2r}, \tag{6.28}$$

where $C_r \leqq 2(8/\pi^2)^{r-1}\Gamma(r+1)$. When combined, Propositions 6.1–6.3, Lemma 6.2, Fact 5 and (6.28), yield readily an alternative proof of Proposition 6.2. In view of Theorem 6.1, we need a refinement of (6.28) in the case $r = 1$. This is provided by the following lemma.

LEMMA 6.3. *Let $Z = Z_{11}$ and $X = X_{11} = W(Z_{11})$ be as in Fact 5. Assume that $H(\beta_1, \beta_2) < \infty$ for some $\beta_1 \geqq 0$ and $\beta_1 \vee \beta_2 > 0$. Then, there exists a constant $B = B_{\beta_1,\beta_2}$ depending upon β_1 and β_2 only such that*

$$\mathbf{E}Z(LZ)^{\beta_1}(LLZ)^{\beta_2} < B\mathbf{E}X^2(L|X|)^{\beta_1}(LL|X|)^{\beta_2} < \infty. \tag{6.29}$$

PROOF. We follow the lines of the Appendix, p. 269–273 of Hall and Heyde [16]. To avoid technicalities, we will limit ourselves to a detailed proof in the case where $\beta_2 = 0$.

Step 1. Let $r = \beta_1$. Let $u > 0$ and $v > 0$ be constants, let $a = (u+v)/2$, $b = (u-v)/(u+v)$, and let Y be a random variable with $\mathbf{E}(Y) = 0$ and $\mathbf{P}(Y = -u \text{ or } v) = 1$. Let $\{W(t),\ t \geqq 0\}$ be a standard Wiener process, and set $T = \inf\{t \geqq 0 : W(t) = -u \text{ or } v\}$. Since the stopping time T on $W(\cdot)$ satisfies $\mathbf{E}W(T) = 0$ and $\mathbf{E}W^2(T) = \mathbf{E}Y^2$, Y and $W(T)$ are identically distributed. We will show that, for each $r \geqq 0$, there exists a constant D_r such that, for all $u, v > 0$,

$$\mathbf{E}T(LT)^r \leqq D_r uv(La)^r. \tag{6.30}$$

Let $\tau = \inf\{t \geqq 0 : W(t) + b = \pm 1\}$. Observe that T and $a^2\tau$ are identically distributed. Moreover, the distribution of τ is given (see e.g. p. 270 in [16]) by

(6.31)
$$\mathbf{P}(\tau > t) = \frac{4}{\pi}\sum_{n=0}^{\infty}(-1)^n\frac{1}{2n+1}\cos\Big(\Big(n+\frac{1}{2}\Big)\pi b\Big)\exp\Big(-\frac{\pi}{8}(2n+1)^2 t\Big) \text{ for } t > 0.$$

We start by proving the following useful inequalities. For any $r \geqq 0$, $a \geqq 0$ and $t \geqq 0$, we have

$$\text{(6.32)}\qquad \begin{aligned} &\text{(i)} && (L(a^2t))^r \leqq (3La)^r(Lt)^r, \quad \text{and} \\ &\text{(ii)} && \frac{d}{dt}(t(Lt)^r) \leqq (1+r)(Lt)^r \quad \text{a.e.}. \end{aligned}$$

Recalling that $Lu = \log(u \vee e)$, we see that for $u' \geqq 0$ and $u'' \geqq 0$, $1 \leqq Lu'u'' \leqq Lu' + Lu''$. Since also $Lt \geqq 1$ and $La^2 \geqq 1$ for all t and a, letting $u' = a^2$ and $u'' = t$ in the above inequality gives

$$(L(a^2t))^r \leqq (Lt)^r(1 + La^2)^r \leqq (Lt)^r(1 + 2La)^r \leqq (Lt)^r(3La)^r,$$

whence (6.32) (i). For $t \leqq e$,

$$\frac{d}{dt}(t(Lt)^r) = 1 \leqq (1 + r)(Lt)^r$$

and, for $t \geqq e$,

$$\frac{d}{dt}(t(Lt)^r) = \left(1 + \frac{r-1}{Lt}\right)(Lt)^r \leqq (1 + r)(Lt)^r,$$

whence (6.32) (ii).

By (6.31) and (6.32), we obtain by integrating by parts that

$$\begin{aligned}
\mathbf{E}T(LT)^r &= a^2\mathbf{E}\tau(La^2\tau)^r \leqq a^2(3La)^r\mathbf{E}\tau(L\tau)^r \\
&= -a^2(3La)^r \int_0^\infty t(Lt)^r d\mathbf{P}(\tau > t) \\
&= a^2(3La)^r \int_0^\infty \frac{d}{dt}(t(Lt)^r)\mathbf{P}(\tau > t)dt \qquad (6.33)\\
&\leqq (1 + r)a^2(3La)^r \int_0^\infty (Lt)^r\mathbf{P}(\tau > t)dt \\
&= (1 + r)a^2(3La)^r g_r(b),
\end{aligned}$$

where, by (6.31),

$$\begin{aligned}
g_r(b) &= \frac{4}{\pi}\sum_{n=0}^\infty (-1)^n \frac{1}{2n+1}\cos\left(\left(n + \frac{1}{2}\right)\pi b\right)\int_0^\infty (Lt)^r \exp\left(-\frac{\pi}{8}(2n+1)^2 t\right) dt \qquad (6.34)\\
&=: \frac{4}{\pi}\sum_{n=0}^\infty (-1)^n \cos\left(\left(n + \frac{1}{2}\right)\pi b\right)\frac{u_n}{(2n+1)^3}.
\end{aligned}$$

The change of variable $z = \frac{\pi^2}{8}(2n+1)^2 t$ in (6.34) shows that

(6.35)
$$0 \leqq u_n = \frac{8}{\pi^2}\int_0^\infty \left(L\left(\frac{8}{\pi^2(2n+1)^2}z\right)\right)^r e^{-z}dz \downarrow \frac{8}{\pi^2}\int_0^\infty e^{-z}dz = \frac{8}{\pi^2} \text{ as } n \uparrow \infty.$$

It follows from (6.34) and (6.35) that $g_r(b)$,

$$g_r'(b) = -2\sum_{n=0}^{\infty}(-1)^n \sin\left(\left(n+\frac{1}{2}\right)\pi b\right)\frac{u_n}{(2n+1)^2}$$

and

$$g_r''(b) = -\sum_{n=0}^{\infty}(-1)^n \cos\left(\left(n+\frac{1}{2}\right)\pi b\right)\frac{u_n}{2n+1}$$

are defined as sums of uniformly convergent series over $b \in [-1,1]$. In the last case, we use Abel's lemma to obtain that, for a suitable constant H_r, $|g_r''(y)| \leqq H_r$ for all $y \in [-1,1]$. Since $g_r(\pm 1) = g_r'(0) = 0$, it follows that, for all $0 \leqq b \leqq 1$,

$$(6.36) \qquad g_r(b) = \int_1^b dx \int_0^x g_r''(y)dy \leqq H_r(1-b^2).$$

A similar argument for $-1 \leqq b \leqq 0$ proves (6.36) for all $b \in [-1,1]$. Since $b = = (u-v)/(u+v)$, $a = (u+v)/2$, and $1-b^2 = uv/a^2$, (6.33) and (6.36) imply (6.30) with $D_r = (1+r)a^2 3^r H_r$.

Step 2. We now consider the case where Y has a continuous distribution function $F(y) = \mathbf{P}(Y \leqq y)$ and satisfies $\mathbf{E}Y = 0$ and $\mathbf{E}Y^2 = 1$. Following the argument of [16] p. 271, based on the original idea of Skorokhod [24], we define a function $G(x) \geqq 0$ for $x \leqq 0$, $G(x) < 0$ for $x > 0$, with

$$\int_x^{G(x)} y dF(y) = 0 \quad \text{for} \quad -\infty < x < \infty.$$

Next, we define the conditional distribution of a random variable X, given $Y = y \neq 0$, by

$$(6.37) \qquad \begin{aligned} \mathbf{P}(X = x \mid Y = y) &= \frac{|G(y)|}{|y| + |G(y)|} \quad \text{and} \\ \mathbf{P}(X = G(y), \mid Y = y) &= \frac{|y|}{|y| + |G(y)|}. \end{aligned}$$

Conditional on $Y = y$, X follows a two-point distribution, and we may, as in Step 1, define a stopping time T on $W(\cdot)$ such that X and $W(T)$ are identically distributed. It follows from the argument, p. 272 of [16], that the unconditioned distribution of X is identical to that of Y, and that the unconditioned distribution of T is identical to that of a random variable Z such that X and $W(Z)$ are identically distributed, with $\mathbf{E}Z = \mathbf{E}X^2 = 1$. To conclude, we make use of the fact that the function $t(Lt)^r$ is increasing and positive on $(0,\infty)$, and hence, satisfies for all $x>0$ and $y>0$ the inequality

$$\left(\frac{x+y}{2}\right)\left(L\left(\frac{x+y}{2}\right)\right)^r \leqq x(Lx)^r + y(Ly)^r.$$

By (6.30) and applying this inequality with $x=|Y|$ and $y=|G(Y)|$, we obtain that

$$\begin{aligned} \mathbf{E}Z(LZ)^r &= \mathbf{E}_Y\left\{\mathbf{E}T(LT)^r|Y)\right\} \\ &\leqq D_r\mathbf{E}_Y\left\{|YG(Y)|\left(L\left(\frac{|Y|+|G(Y)|}{2}\right)\right)^r\right\} \\ &\leqq 2D_r\mathbf{E}_Y\left\{|YG(Y)|(L|Y|)^r\frac{|Y|}{|Y|+|G(Y)|}\right. \\ &\qquad \left. + |YG(Y)|(L|G(Y)|)^r\frac{|G(Y)|}{|Y|+|G(Y)|}\right\} \\ &= 2D_r\mathbf{E}_Y\{\mathbf{E}(X^2(L|X|)^r|X=Y)\mathbf{P}(X=Y) \\ &\qquad + \mathbf{E}(X^2(L|X|)^r|X=G(Y))\mathbf{P}(X=G(Y))\} \\ &= 2D_r\mathbf{E}X^2(L|X|)^r. \end{aligned} \tag{6.38}$$

Thus, by (6.38), we obtain (6.29) by setting $B=2D_r$.

Step 3. Having established the lemma when Y has a continuous distribution function, we proceed as on p. 272 of [16], using the fact that an arbitrary distribution may be suitably approximated by a continuous distribution, to conclude that the result holds in general. □

Proof of Theorem 6.1. It follows directly from Propositions 6.1–6.3, Lemmas 6.2–6.3, and Fact 5. □

Remark 6.1. Theorem 6.1 and Proposition 2.1 show that the condition $H(\beta_1,\beta_2)<\infty$ (resp. $H(\beta_1,\beta_2-1)=\infty$) is sufficient to ensure that the conclusion of Theorem 1.2 holds (resp. does not hold) for $\mathcal{X}$, $-\mathcal{X}$ and $k_n = (Ln)^{\beta_1}(LLn)^{\beta_2}$. This illustrates the sharpness of our results.

Remark 6.2. The Skorohod embedding may be used as above for non-identically distributed random variables with a martingale structure (see e.g. [16]). Laws of the iterated logarithm for row-sums of arrays of this type will be considered elsewhere.

Acknowledgement. We thank the referee for a careful reading of our manuscript and for valuable comments.

REFERENCES

[1] Baxter, G., An analogue of the law of the iterated logarithm, *Proc. Amer. Math. Soc.* **6** (1955), 177–181. *MR* **16**-1128

[2] Chow, J. S. and Teicher, H., *Probability theory. Independence, interchangeability, martingales*, Springer-Verlag, New York – Heidelberg, 1978. *MR* **80a**:60004

[3] Cramér, H., Su un teorema relativo alla legge uniforme dei grandi numeri, *Giorn. Ist. Ital. Attuari* **5** (1934), 1–13. *Zbl.* **8**, 368

[4] Csörgő, M. and Révész, P., *Strong approximations in probability and statistics*, Probability and Mathematical Statistics, Academic Press, New York – London, 1981. *MR* **84d**:60050

[5] Csörgő, S. and Hall, P., Upper and lower classes for triangular arrays, *Z. Wahrsch. Verw. Gebiete* **61** (1982), 207–222. *MR* **84f**:60044

[6] Dabrowski, A., Dehling, H. and Philipp, W., An almost sure invariance principle for triangular arrays of Banach space valued random variables, *Z. Wahrsch. Verw. Gebiete* **65** (1984), 483–491. *MR* **85h**:60008

[7] Deheuvels, P. and Steinebach, J., Strong laws for extreme values of sequences of partial sum processes (submitted).

[8] Deheuvels, P. and Révész, P., On the coverage of Strassen-type sets by sequences of Wiener processes, *J. Theoret. Probab.* **6** (1993), 427–449. *MR* **94k**:60063

[9] Eicker, F., A log log-law for double sequences of random variables, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **16** (1970), 107–133. *MR* **43** #4091

[10] Einmahl, U. and Mason, D. M., Some results on the almost sure behaviour of martingales, *Limit theorems in probability and statistics* (Pécs, 1989), Colloq. Math. Soc. János Bolyai, **57**, North-Holland, Amsterdam, 1990, 185–195. *MR* **92g**:60041

[11] Esseen, C.-G., Fourier analysis of distribution functions. A mathematical study of the Laplace–Gaussian law, *Acta. Math.* **77** (1945), 1–125. *MR* **7**–312

[12] Feller, W., A limit theorem for random variables with infinite moments. *Amer. J. Math.* **68** (1946), 257–262. *MR* **8**–37

[13] Fernholz, L. T., Limit points associated with generalized iterated logarithm laws, *Ann. Probab.* **8** (1980), 390–392. *MR* **81f**:60045

[14] Finkelstein, H., The law of the iterated logarithm for empirical distributions, *Ann. Math. Statist.* **42** (1971), 607–615. *MR* **44** #4803

[15] Gut, A., The weak law of large numbers for arrays, *Statist. Probab. Lett.* **14** (1992), 49–52. *MR* **93a**:60028

[16] Hall, P. and Heyde, C. C., *Martingale limit theory and its application*, Probability and Mathematical Statistics, Academic Press, New York – London, 1980. *MR* **83a**:60001

[17] Hartman, P. and Wintner, A., On the law of the iterated logarithm, *Amer. J. Math.* **63** (1941), 169–176. *MR* **2**–228

[18] Komlós, J., Major, P. and Tusnády, G., An approximation of partial sums of independent RV's and the sample DF. I, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **32** (1975), 111–131. *MR* **51** #11605b

[19] Komlós, J., Major, P. and Tusnády, G., An approximation of partial sums of independent RV's and the sample DF. II, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **34** (1976), 33–58. *MR* **53** #6697

[20] Lai, T. L. and Wei, C. Z., A law of the iterated logarithm for double arrays of independent random variables with applications to regression and time series models, *Ann. Probab.* **10** (1982), 320–335. *MR* **83h**:60029

[21] Le Page, R. and Schreiber, B. M., An iterated logarithm law for families of Brownian paths, *Z. Wahrsch. Verw. Gebiete* **70** (1985), 341–344. *MR* **86m**:60198

[22] Major, P., On the invariance principle for sums of independent identically distributed random variables, *J. Multivariate Anal.* **8** (1978), 487–517. *MR* **80f**:60034

[23] Shorack, G. R. and Wellner, J. A., *Empirical processes with applications to statistics*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, Wiley, New York, 1986. *MR* **88e**:60002

[24] SKOROHOD, A. V., *Studies in the theory of random processes*, Addison-Wesley, Reading, MA, 1965. *MR* **32** # 3082b

[25] SMYTHE, R. T., Strong laws of large numbers for r-dimensional arrays of random variables, *Ann. Probab.* **1** (1973), 164–170. *MR* **49** #11602

[26] STRASSEN, V., An invariance principle for the law of the iterated logarithm, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **3** (1964), 211–226. *MR* **30** #5379

[27] TEICHER, H., On the law of the iterated logarithm, *Ann. Probab.* **2** (1974), 714–728. *MR* **51** #6941

[28] TEICHER, H., Generalized exponential bounds, iterated logarithm and strong laws, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **48** (1979), 293–307. *MR* **80h**:60044

[29] TEICHER, H., Almost certain behaviour of row sums of double arrays, *Analytical methods in probability theory* (Oberwolfach, 1980), Lecture Notes in Mathematics, 861, Springer-Verlag, Berlin – New York, 1981, 155–165. *MR* **83f**:60055

[30] TEICHER, H., Almost certain convergence in double arrays, *Z. Wahrsch. Verw. Gebiete* **69** (1985), 331–345. *MR* **86h**:60071

[31] TOMKINS, R., On the law of the iterated logarithm for double sequences of random variables, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **30** (1974), 303–314. *MR* **53** #9348

(Received July 25, 1994)

UNIVERSITÉ PIERRE ET MARIE CURIE
LABORATOIRE DE STATISTIQUE
THÉORIQUE ET APPLIQUÉE
TOUR 45–55-3^e–BOÎTE NO. 158
4, PLACE JUSSIEU
F–75252 PARIS Cedex 05
FRANCE

DEPARTMENT OF STATISTICS
HILL CENTER
BUSH CAMPUS
RUTGERS UNIVERSITY
NEW BRUNSWICK, NJ 08903
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 187–196

A STRONG APPROXIMATION FOR LOGARITHMIC AVERAGES

L. HORVÁTH and D. KHOSHNEVISAN

Dedicated to P. Révész for his sixtieth birthday

Summary

Let $S_1, S_2, \dots$ be a sequence of partial sums of random variables. We obtain limit theorems for $\sum_{1 \leqq k \leqq n} \frac{1}{k} f(S_k/k^{1/2})$.

1. Introduction and results

Let $X, X_1, X_2, \dots$ be a sequence of independent, identically distributed random variables with

C1 $$\mathbf{E}X = 0 \quad \text{and} \quad \mathbf{E}X^2 = 1.$$

Brosamler [2], Schatte [13]–[15], Lacey and Philipp [10], Berkes and Dehling [1], Weigl [16], Csörgő and Horváth [4] and Csáki, Földes and Révész [3] investigated the asymptotic properties of the logarithmic averages

$$T_n = \frac{1}{\log n} \sum_{1 \leqq k \leqq n} \frac{1}{k} f(S_k/k^{1/2}),$$

where

(1.1) $$S_t = \sum_{1 \leqq i \leqq t} X_i, \qquad 1 \leqq t < \infty.$$

Schatte [14] showed that

(1.2) $$\lim_{n \to \infty} T_n = \int_{-\infty}^{\infty} f(x)\varphi(x)dx \qquad \text{a.s.,}$$

1991 *Mathematics Subject Classification.* Primary 60F15, 60F17; Secondary 60J15.
Key words and phrases. Strong approximation, logarithmic average, Ornstein–Uhlenbeck process, partial sums, laws of the iterated logarithm.

[1] Research supported by an NSF–NATO East-Europe Outreach Grant.

where $\varphi(x) = (2\pi)^{-1/2}e^{-x^2/2}$ is the standard normal density function, if $|f(x)| \leqq e^{\gamma x^2}$ with some $\gamma < 1/4$ and $\mathbf{E}|X|^{2+\delta} < \infty$ with some $\delta > 0$. If $f(x) = I\{x \leqq x_0\}$, then the almost sure limit is $\Phi(x_0)$ (Φ stands for the standard normal distribution function), and in this case (1.2) is called an almost sure central limit theorem. Weigl [16] (cf. Révész [12]) and Csörgő and Horváth [4] showed that

$$(\log n)^{-1/2}\left\{\sum_{1\leqq k\leqq n}\frac{1}{k}\left(I\left\{\frac{S_k}{k^{1/2}}\leqq x_0\right\}-\Phi(x_0)\right)\right\}\xrightarrow{\mathcal{D}}\sigma_0 N(0,1), \tag{1.3}$$

where σ_0 is a constant, depending only on x_0, and $N(0,1)$ is a standard normal random variable. Csáki, Földes and Révész [3] asked whether (1.3) remains true for a larger class of functions f. In this paper we obtain approximations for

$$A(n) = \sum_{1\leqq k\leqq n}\frac{1}{k}f(S_k/k^{1/2}) - \log n\int_{-\infty}^{\infty} f(x)\varphi(x)dx,$$

which will also imply the asymptotic normality of $(\log n)^{-1/2}A(n)$.

We assume that f satisfies the following conditions:

C2 $f(x) = f_1(x) - f_2(x)$, where f_1 and f_2 are non-decreasing functions;

C3 there is some $0 < \alpha \leqq 1$ such that

$$\int_{-\infty}^{\infty}|f_1(x+t)-f_1(x)|e^{-x^2/2}dx = O(|t|^\alpha), \quad \text{as} \quad t\to 0 \tag{1.4}$$

and

$$\int_{-\infty}^{\infty}|f_2(x+t)-f_2(x)|e^{-x^2/2}dx = O(|t|^\alpha), \quad \text{as} \quad t\to 0 \tag{1.5}$$

and

C4 $$\int_{-\infty}^{\infty}(|f_1(x)|+|f_2(x)|)^{3+\delta}e^{-x^2/2}dx<\infty \qquad \text{with some } \delta>0.$$

Let

$$\begin{aligned}\sigma^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x)f(y)\Bigg\{&\frac{1}{2\pi(1-e^{-|u|})^{1/2}}\times\\ \times\exp\left(-\frac{1}{2(1-e^{-|u|})}(x^2+y^2-2e^{-|u|/2}xy)\right)&-\varphi(x)\varphi(y)\Bigg\}dxdydu.\end{aligned} \tag{1.6}$$

We note that $\sigma^2 < \infty$ by C4. Our first result is a strong approximation for $A(n)$. We need more than two moments:

C5 $$\mathbf{E}|X|^2(\log|X|)^{\delta+2/\alpha} < \infty \quad \text{for some } \delta > 0.$$

THEOREM 1.1. *If C1–C5 hold, then we can find a Wiener process $\{W(t), 0 \leqq t < \infty\}$ such that*

$$|A(n) - \sigma W(\log n)| \overset{a.s.}{=} o((\log n)^{1/3}).$$

The strong approximation in Theorem 1.1 gives immediately the weak convergence of $(\log n)^{-1/2}A(n^t)$ as well as the laws of the iterated logarithm for $A(n)$.

COROLLARY 1.1. *If C1–C5 hold, then we have*

$$(\sigma^2 \log n)^{-1/2} A(n^t) \xrightarrow{\mathcal{D}[0,1]} W(t),$$

where $\{W(t), 0 \leqq t \leqq 1\}$ is a Wiener process,

$$\limsup_{n\to\infty} (\sigma^2 \log n)^{-1/2} (2 \log\log\log n)^{-1/2} \max_{1\leqq k \leqq n} |A(k)| = 1 \quad a.s.$$

and

$$\liminf_{n\to\infty} (\sigma^2 \log n)^{-1/2} (\log\log\log n)^{1/2} \max_{1\leqq k\leqq n} |A(k)| = \frac{\pi}{8^{1/2}} \quad a.s.$$

Corollary 1.1 follows immediately from the properties of the Wiener process (cf. Csörgő and Révész [5]) and Theorem 1.1. Next we discuss a few examples which were considered earlier in the literature.

EXAMPLE 1.1. Let $f(x) = I\{x \leq x_0\}$. Now C2 holds with $f = f_1$ and $f_2 = 0$. Elementary arguments give that,

$$\int_{-\infty}^{\infty} |f(x+t) - f(x)|\varphi(x)dx = |\Phi(t + x_0) - \Phi(x_0)|.$$

Since Φ has a bounded derivative, C3 holds with $\alpha = 1$. We have C4 for all $\delta > 0$, because f is bounded. Csörgő and Horváth [4] proved earlier Corollary 1.1 in this special case.

EXAMPLE 1.2. Let $f(x) = |x|^\gamma$, $\gamma \geqq 1$. It is clear that C2 is satisfied with $f_1(x) = x^\gamma I\{x \geqq 0\}$ and $f_2(x) = -(-x)^\gamma I\{x < 0\}$ and C4 holds for all $\gamma > 0$. Let N be a standard normal random variable. If $t > 0$, then we have

$$\begin{aligned}\mathbf{E}|(N+t)^\gamma I\{N+t \geqq 0\} - N^\gamma I\{N \geqq 0\}| &\leqq \mathbf{E}\gamma(|N|+t)^{\gamma-1}t \\ &\quad + \mathbf{E}(|N|+t)^\gamma I\{-t < N < 0\}.\end{aligned}$$

By Hölder's inequality we get

$$\mathbf{E}(|N|+t)^{\gamma} I\{-t<N<0\}=O(t^{\alpha})$$

for all $0<\alpha<1$. Similarly, if $t<0$, then

$$\begin{aligned}
&\mathbf{E}|(N+t)^{\gamma} I\{N+t \geqq 0\}-N^{\gamma} I\{N \geqq 0\}| \\
&\leqq \mathbf{E}\gamma(|N|+|t|)^{\gamma-1}|t|+\mathbf{E}(|N|+|t|)^{\gamma} I\{0<N<-t\} \\
&=O(|t|^{\alpha})
\end{aligned}$$

for all $0<\alpha<1$. Thus (1.4) holds with any $0<\alpha<1$ and similar arguments give (1.5).

EXAMPLE 1.3. Let $f(x)=e^{\gamma|x|}$, $0<\gamma<\infty$. We have C2 with $f_1(x)=e^{\gamma x} I\{x \geqq 0\}$ and $f_2(x)=-e^{-\gamma x} I\{x \leqq 0\}$ and C4 is satisfied with any $\delta>0$. Let N be a standard normal random variable again. If $t \geqq 0$, then we have by Hölder's inequality

$$\begin{aligned}
\mathbf{E}|f_1(N+t)-f_1(N)| &\leqq \mathbf{E}|e^{\gamma(N+t)}-e^{\gamma N}| I\{N \geqq 0\} \\
&\quad+\mathbf{E}e^{\gamma(|N|+t)} I\{-t \leqq N \leqq 0\} \\
&=O(t^{\alpha})
\end{aligned}$$

for all $0<\alpha<1$. Similarly, if $t<0$, then

$$\mathbf{E}|f_1(N+t)-f_1(N)|=O(|t|^{\alpha})$$

for all $0<\alpha<1$, which yields (1.4). Since $-N$ is also standard normal, (1.4) implies (1.5).

EXAMPLE 1.4. Let $f(x)=e^{\gamma x^2}$, $0<\gamma<1/6$. Then C2 is satisfied with $f_1(x)=e^{\gamma x^2} I\{x \geqq 0\}$ and $f_2(x)=-e^{\gamma x^2} I\{x<0\}$ and C4 holds with $0<\delta<\frac{1}{2\gamma}-3$. As in Examples 1.2 and 3 we get

$$\begin{aligned}
\mathbf{E}|f_i(N+t)-f_i(N)| &\leqq |t| \mathbf{E} e^{\gamma(|N|+|t|)^2}(2\gamma|N|+|t|) \\
&\quad+2\mathbf{E}e^{\gamma(|N|+|t|)^2} I\{0<N<|t|\},
\end{aligned}$$

$i=1,2$. By Hölder's inequality we have

$$\begin{aligned}
&\mathbf{E}e^{\gamma(|N|+|t|)^2} I\{0<N<|t|\} \\
&\leqq (\mathbf{E}e^{\beta(|N|+|t|)^2})^{\gamma/\beta}(\Phi(|t|)-\Phi(0))^{(\beta-\gamma)\beta} \\
&=O(|t|^{(\beta-\gamma)/\beta})
\end{aligned}$$

for all $0<\beta<1/2$. Hence C3 holds with any $0<\alpha<1-2\gamma$.

2. Proofs

The proof of Theorem 1.1 is based on two lemmas. The first lemma shows that it is enough to consider the case when X is a standard normal random variable. We prove an almost sure approximation for integrals of Ornstein–Uhlenbeck processes in the second lemma.

LEMMA 2.1. *If C1–C3 and C5 hold, then we can find a Wiener process* $\{W^*(t),\ 0 \leqq t < \infty\}$ *such that*

$$\left| A(n) + \log n \int_{-\infty}^{\infty} f(x)\varphi(x)dx - \int_{1}^{n} \frac{1}{t} f(W^*(t)/t^{1/2})dt \right| \overset{a.s.}{=} o(1).$$

PROOF. By Einmahl [6] we can define a Wiener process $\{W^*(t), 0 \leqq t < \infty\}$ such that

$$S_t - W^*(t) \overset{a.s.}{=} o(t^{1/2}(\log t)^{-\delta/2-1/\alpha}), \quad \text{as} \quad t \to \infty.$$

Hence we can find a random variable t_0 such that

$$\left| \frac{S_t}{[t]^{1/2}} - \frac{W^*(t)}{t^{1/2}} \right| \leqq (\log t)^{-\delta/2-1/\alpha},$$

if $t \geqq t_0$, where $[t]$ is the integer part of t. Since f_1 and f_2 are a non-decreasing for $i = 1, 2$ we have

$$\begin{aligned} \int_{t_0}^{t} \frac{1}{u} f_i\left(\frac{W^*(u)}{u^{1/2}} - (\log u)^{-\delta/2-1/\alpha}\right) du &\leqq \int_{t_0}^{t} \frac{1}{u} f_i\left(\frac{S_u}{[u]^{1/2}}\right) du \\ &\leqq \int_{t_0}^{t} \frac{1}{u} f_i\left(\frac{W^*(u)}{u^{1/2}} + (\log u)^{-\delta/2-1/\alpha}\right) du, \end{aligned} \tag{2.1}$$

if $t \geqq t_0$. It follows from C3 that

$$\int_{3}^{\infty} \frac{1}{u} \left| f_i\left(\frac{W^*(u)}{u^{1/2}} + (\log u)^{-\delta/2-1/\alpha}\right) - f_i\left(\frac{W^*(u)}{u^{1/2}}\right) \right| du < \infty \quad \text{a.s.} \tag{2.2}$$

and

$$\int_{3}^{\infty} \frac{1}{u} \left| f_i\left(\frac{W^*(u)}{u^{1/2}} - (\log u)^{-\delta/2-1/\alpha}\right) - f_i\left(\frac{W^*(u)}{u^{1/2}}\right) \right| du < \infty \quad \text{a.s.} \tag{2.3}$$

Putting together (2.1) – (2.3) we get that

$$\sup_{t_0 \leqq t < \infty} \left| \int_{t_0}^{t} \frac{1}{u} f\left(\frac{S_u}{[u]^{1/2}}\right) du - \int_{t_0}^{t} \frac{1}{u} f\left(\frac{W^*(u)}{u^{1/2}}\right) du \right| < \infty \quad \text{a.s.,}$$

and therefore we have

$$\sup_{1 \leqq t < \infty} \left| \int_{1}^{t} \frac{1}{u} f\left(\frac{S_u}{[u]^{1/2}}\right) du - \int_{1}^{t} \frac{1}{u} f\left(\frac{W^*(u)}{u^{1/2}}\right) du \right| < \infty \quad \text{a.s.}.$$

It is easy to see that

$$\begin{aligned} \sup_{1 \leqq n < \infty} & \left| A(n) + \log n \int_{-\infty}^{\infty} f(x)\varphi(x)dx - \int_{1}^{n} \frac{1}{u} f\left(\frac{S_u}{[u]^{1/2}}\right) du \right| \\ & \leqq C \sum_{1 \leqq n < \infty} \frac{1}{n^2} \left| f\left(\frac{S_n}{n^{1/2}}\right) \right|, \\ & \leqq 2C \int_{1}^{\infty} \frac{1}{u^2} \left| f\left(\frac{S_u}{[u]^{1/2}}\right) \right| du \end{aligned}$$

with some constant C. Similarly to (2.2) and (2.3) we have

$$\begin{aligned} \int_{1}^{\infty} \frac{1}{u^2} \left| f\left(\frac{S_u}{[u]^{1/2}}\right) \right| du \leqq & \sup_{1 \leqq t < \infty} \left| \int_{1}^{t} \frac{1}{u^2} \left(f_1\left(\frac{S_u}{[u]^{1/2}}\right) - f_1\left(\frac{W^*(u)}{u^{1/2}}\right) \right) du \right| \\ & + \sup_{1 \leqq t < \infty} \left| \int_{1}^{t} \frac{1}{u^2} \left(f_2\left(\frac{S_u}{[u]^{1/2}}\right) - f_2\left(\frac{W^*(u)}{u^{1/2}}\right) \right) du \right| \\ & + \int_{1}^{\infty} \frac{1}{u^2} \left\{ \left| f_1\left(\frac{W^*(u)}{u^{1/2}}\right) \right| + \left| f_2\left(\frac{W^*(u)}{u^{1/2}}\right) \right| \right\} du \\ & < \infty \quad \text{a.s.,} \end{aligned}$$

which completes the proof of Lemma 2.1.

Let $\{V(t),\ -\infty < t < \infty\}$ be an Ornstein–Uhlenbeck process. This means that $V(t)$ is a stationary Gaussian process with $\mathbf{E}V(t) = 0$ and $\mathbf{E}V(t)V(s) = \exp(-|t-s|/2)$.

LEMMA 2.2. *If C4 holds, then we can find a Wiener process* $\{W(t), 0 \leqq t < \infty\}$ *such that*

$$\left|\int_0^t \{f(V(u)) - \int_{-\infty}^{\infty} f(s)\varphi(s)ds\}du - \sigma W(t)\right| \overset{\text{a.s.}}{=} o(t^{1/3}), \quad as \quad t \to \infty,$$

where σ *is defined in* (1.6).

PROOF. Let $\bar{f}(x) = f(x) - \int_{-\infty}^{\infty} f(t)\varphi(t)dt$ and define

$$\tau_{2j+1} = \inf\{s : s > \tau_{2j}, V(s) = 1\}$$

and

$$\tau_{2j+2} = \inf\{s : s > \tau_{2j+1}, V(s) = 0\}.$$

We note that $\{\tau_{2j+2} - \tau_{2j}, 1 \leqq j < \infty\}$ are independent identically distributed random variables with a finite moment generating function at zero (cf. Lemma 2.3 in Horváth and Khoshnevisan [7]). Let

$$\xi_j = \int_{\tau_{2j}}^{\tau_{2j+2}} \bar{f}(V(t))dt.$$

It is easy to see that $\{\xi_j, 1 \leqq j < \infty\}$ are independent identically distributed random variables. We show that

$$\mathbf{E}|\xi_1|^3 < \infty. \tag{2.3}$$

Let $X(t) = |\bar{f}(V(t))| I\{\tau_2 \leqq t \leqq \tau_4\}$. Then

$$\mathbf{E}|\xi_1|^3 \leqq \mathbf{E}(\int_0^{\infty} X(t)dt)^3$$
$$= \int_0^{\infty}\int_0^{\infty}\int_0^{\infty} \mathbf{E}X(t)X(u)X(v)dtdudv.$$

Using Hölder's inequality we get

$$\mathbf{E}X(t)X(u)X(v) \leqq (\mathbf{E}X^3(t))^{1/3}(\mathbf{E}X^3(u))^{1/3}(\mathbf{E}X^3(v))^{1/3},$$

and therefore

$$\mathbf{E}|\xi_1|^3 \leqq \left(\int_0^{\infty} (\mathbf{E}X^3(t))^{1/3}dt\right)^3. \tag{2.5}$$

Also,

$$\mathbf{E}X^3(t) \leqq (\mathbf{E}|\bar{f}(V(t))|^{3+\delta})^{\frac{3}{3+\delta}}(\mathbf{E}(I\{\tau_2 \leqq t \leqq \tau_4\})^{\frac{3+\delta}{\delta}})^{\frac{\delta}{3+\delta}}$$

$$= (\mathbf{E}|\bar{f}(V(t))|^{3+\delta})^{\frac{3}{3+\delta}}(\mathbf{P}\{\tau_2 \leqq t \leqq \tau_4\})^{\frac{\delta}{3+\delta}}. \tag{2.6}$$

Since $V(t)$ is a standard normal random variable for each t, condition C4 implies that

$$\mathbf{E}|\bar{f}(V(t))|^{3+\delta} = \mathbf{E}|\bar{f}(V(1))|^{3+\delta} < \infty. \tag{2.7}$$

The random variables τ_2 and $\tau_4 - \tau_2$ have finite moment generating functions in a neighbourhood of zero and therefore, we can find two positive constants α and β such that

$$\mathbf{P}\{\tau_2 \leqq t \leqq \tau_4\} \leqq \beta e^{-\alpha t}. \tag{2.8}$$

Now (2.4) follows from (2.5)–(2.8).

We note that $\mathbf{E}\xi_i = 0$ (cf. Mandl [11], p.95) and $\tilde{\sigma}^2 = \operatorname{var}\xi_i < \infty$ by (2.4). Using the Komlós, Major and Tusnády [8], [9] approximation we can find a Wiener process $\{\tilde{W}(t), 0 \leqq t < \infty\}$ such that

$$\sum_{1 \leqq j \leqq x} \xi_j - \tilde{\sigma}\tilde{W}(x) \overset{a.s.}{=} o(x^{1/3}), \quad \text{as} \quad x \to \infty. \tag{2.9}$$

Let $m = \mathbf{E}(\tau_4 - \tau_2)$. The law of the iterated logarithm for partial sums yields

$$|\tau_{2k} - km| \overset{a.s.}{=} O((k \log\log k)^{1/2}), \quad \text{as} \quad k \to \infty. \tag{2.10}$$

It follows from (2.6)–(2.8) that

$$\mathbf{E}\left(\int_{\tau_2}^{\tau_4} |\bar{f}(V(t))| dt\right)^3 < \infty$$

and therefore

$$\int_{\tau_{2k}}^{\tau_{2k+2}} |\bar{f}(V(t))| dt \overset{a.s.}{=} o(k^{1/3}), \quad \text{as} \quad k \to \infty. \tag{2.11}$$

By (2.11) we have that

$$\max_{\tau_{2k} \leqq T \leqq \tau_{2k+2}} \left| \int_0^T \bar{f}(V(t)) dt - \int_0^{\tau_{2k}} \bar{f}(V(t)) dt \right|$$

$$\leqq \int_{\tau_{2k}}^{\tau_{2k+2}} |\bar{f}(V(t))| dt$$

$$\overset{a.s.}{=} o(k^{1/3}), \quad \text{as} \quad k \to \infty. \tag{2.12}$$

The modulus of continuity of the Wiener process (cf. Csörgő and Révész [5]), (2.10) and (2.12) gives

$$\begin{aligned}\left|\int_0^T f(V(t))dt - \tilde{\sigma}\bar{W}(T/m)\right| &\leq \left|\int_0^{\tau_{2k}} f(V(t))dt - \tilde{\sigma}\bar{W}(k)\right| \\ &+ \left|\int_0^T f(V(t))dt - \int_0^{\tau_{2k}} f(V(t))dt\right| \\ &+ |\tilde{\sigma}\bar{W}(k) - \tilde{\sigma}\bar{W}(T/m)| \\ &\overset{\text{a.s.}}{=} o(T^{1/3}) + O((T\log\log T)^{1/4}(\log T)^{1/2}),\end{aligned} \tag{2.13}$$

where k is defined by $\tau_{2k} \leq T < \tau_{2k+2}$. It is easy to check that $W(x) = m^{1/2}\bar{W}(x/m)$ is a Wiener process and therefore Lemma 2.2 is established with $\sigma^2 = \tilde{\sigma}^2/m$.

Next we show that σ^2 can be written as the integral in (1.6). Replacing the almost sure arguments with estimates for the second moments we get (cf. also Mandl [11]) that

$$\lim_{T\to\infty} \frac{1}{T} \operatorname{var}\left(\int_0^T f(V(t))dt\right) = \tilde{\sigma}^2/m.$$

Elementary calculations give that

$$\lim_{T\to\infty} \frac{1}{T}\mathbf{E}\left\{\int_0^T f(V(t))dt\right\}^2 =$$

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T\int_0^T \mathbf{E}f(V(t))f(V(s))dsdt = \sigma^2,$$

where σ^2 is defined in (1.6). Hence the proof of Lemma 2.2 is complete.

Proof of Theorem 1.1. By Lemma 2.1 it is enough to get strong approximation for $\int_1^n \frac{1}{t} f(W^*(t)/t^{1/2})dt$. It is easy to see that $W^*(t)/t^{1/2} = V(\log t)$, $1 \leq t < \infty$, where $\{V(x), 0 \leq x < \infty\}$ is an Ornstein–Uhlenbeck process. Since

$$\int_1^n \frac{1}{t} f(V(\log t))dt = \int_0^{\log n} f(V(u))du,$$

Lemma 2.2 implies Theorem 1.1

REFERENCES

[1] BERKES, I. and DEHLING, H., Some limit theorems in log density, *Ann. Probab.* **21** (1993), 1640–1670. *MR* **94h**:60026

[2] BROSAMLER, G., An almost everywhere central limit theorem, *Math. Proc. Cambridge Philos. Soc.* **104** (1988), 561–574. *MR* **89i**:60045

[3] CSÁKI, E., FÖLDES, A. and RÉVÉSZ, P., On almost sure local and global central limit theorems, *Probab. Theory Related Fields* **97** (1992), 521–557.

[4] CSÖRGŐ, M. and HORVÁTH, L., Invariance principles for logarithmic averages, *Math. Proc. Cambridge Philos. Soc.* **112** (1992), 195–205. *MR* **93e**:60057

[5] CSÖRGŐ, M. and RÉVÉSZ, P., *Strong approximations in probability and statistics*, Probability and Mathematical Statistics, Academic Press, New York, 1981. *MR* **84d**:60050

[6] EINMAHL, U., Strong invariance principles for partial sums of independent random vectors, *Ann. Probab.* **15** (1987), 1419–1440. *MR* **88h**:60071

[7] HORVÁTH, L. and KHOSHNEVISAN, D., Weight functions and pathwise local central limit theorems, *Stochastic Processes Appl.* (1995) (to appear).

[8] KOMLÓS, J., MAJOR, P. and TUSNÁDY, G., An approximation of partial sums of independent R.V.'s and the sample DF.I, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **32** (1975), 111–131. *MR* **51** # 11605b

[9] KOMLÓS, J., MAJOR, P. and TUSNÁDY, G., An approximation of partial sums of independent R.V.'s and the sample DF.II, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **34** (1976), 33–58. *MR* **53** # 6697

[10] LACEY, M. and PHILIPP, W., A note on the almost sure central limit theorem, *Statist. Probab. Lett.* **9** (1990), 201–205. *MR* **91e**:60100

[11] MANDL, P., *Analytical treatment of one-dimensional Markov processes*, Die Grundlehren der math. Wissenschaften, Bd. 151, Academia, Prague; Springer-Verlag, New York, 1968. *MR* **40** # 930

[12] RÉVÉSZ, P., *Random walk in random and non-random environments*, World Scientific, Singapore, 1990. *MR* **92c**:60096

[13] SCHATTE, P., On strong versions of the central limit theorem, *Math. Nachr.* **137** (1988), 249–256. *MR* **89i**:60070

[14] SCHATTE, P., On the central limit theorem with almost sure convergence, *Probab. Math. Statist.* **11** (1990), 237–246. *MR* **92k**:60050

[15] SCHATTE, P., Two remarks on the almost sure central limit theorem, *Math. Nachr.* **154** (1991), 225–229. *MR* **93b**:60066

[16] WEIGL, A., *Zwei Sätze über die Belegungszeit beim Random Walk*, Diplomarbeit, Technische Universität, Wien, 1989.

(Received May 20, 1994)

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF UTAH
SALT LAKE CITY, UT 84112
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 197–202

ON THE ALMOST SURE CENTRAL LIMIT THEOREM FOR Φ–MIXING RANDOM VARIABLES

G. HURELBAATAR

Dedicated to P. Révész for his sixtieth birthday

1. Introduction

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, $\{X_i, i \geqq 1\}$ a sequence of r.v.'s and let M_a^b be the σ-algebra generated by the random variables $\{X_i, a \leqq i \leqq b\}$. Define

$$\phi(n) = \sup_{k \geqq 1} \sup_{A \in M_1^k, B \in M_{k+n}^{\infty}} |\mathbf{P}(B|A) - \mathbf{P}(B)|.$$

The sequence $\{X_i, i \geqq 1\}$ is said to be ϕ–mixing if $\phi(n) \to 0$ as $n \to \infty$.

Setting $S_n = X_1 + \cdots + X_n$ define the usual "*broken line*" process $\hat{s}_n(t)$ on [0,1] for some numerical sequence $a_n > 0$, by

$$\hat{s}_n(t) = \begin{cases} \dfrac{S_j}{a_n}, & \text{for } t = t_{nj}, \quad 0 \leqq j \leqq n \\ \text{linear in between.} \end{cases} \tag{1}$$

Here for every $n \geqq 1$, $0 = t_{n0} < t_{n1} < \ldots < t_{nn} = 1$ is a partition of [0,1] such that

$$\max_{1 \leqq j \leqq n} |t_{n,j-1} - t_{n,j}| \to 0 \quad \text{as} \quad n \to \infty.$$

If $t_{n_j} = \frac{j}{n}$, then we call it uniform partition. In the sequel we shall denote by $C, C_1, C_2, \ldots$ positive constants and the weak convergence by $\Longrightarrow$.

In this paper we prove the following theorem.

THEOREM 1. *Let $X_1, X_2, \ldots$ be a sequence of ϕ–mixing r.v.'s, with $\mathbf{E}X_i = 0$, satisfying*

$$\sum_{n=1}^{\infty} \phi^{\frac{1}{2}}(2^n) < \infty \tag{2}$$

1991 *Mathematics Subject Classification*. Primary 60F15, 60F05.

Key words and phrases. Almost sure central limit theorem, logarithmic average, weakly dependent random variables.

Research supported by Hungarian National Foundation for Scientific Research Grant No. 1905.

$$\sum_{i=1}^{n} \mathbf{E}X_i^2 < Ca_n^2 \qquad (n \geqq 1), \tag{3}$$

where a_n is the same as in (1). Assume that

$$\frac{a_l}{a_k} \geqq c\left(\frac{l}{k}\right)^{\gamma} \qquad (l \geqq k) \tag{4}$$

with some $\gamma > 0$. Then for any Borel-subset A of $D[0,1]$ with $\mu(\partial A) = 0$ (μ is a probability measure on $D[0,1]$) the following two statements are equivalent:

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} I\{\hat{s}_k(\cdot) \in A\} = \mu(A) \qquad a.s. \tag{5}$$

and the exceptional set of probability zero does not depend on A. Here I denotes the indicator function.

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} P\{\hat{s}_k(\cdot) \in A\} = \mu(A). \tag{6}$$

The equivalence of the above two statements was proved by Berkes and Dehling [2] for independent random variables under the condition

$$\mathbf{E}f\left(\left|\frac{S_n - b_n}{a_n}\right|\right) \leqq (\log\log n)^{-1-\varepsilon} f(e^{(\log n)^{1-\varepsilon}}) \qquad (n \geqq n_0)$$

for some function f and numerical sequences $a_n > 0$ and b_n.

Corollary. *Let $\{X_i, i \geqq 1\}$ be a sequence of ϕ-mixing random variables, with $\mathbf{E}X_i = 0$, $\mathbf{E}|X_i|^{2+\varepsilon} < C\,(i \geqq 1)$ satisfying (2) and*

$$\lim_{n\to\infty} \frac{\mathbf{E}S_n^2}{n} = \sigma^2 > 0.$$

Let $a_n = \sigma\sqrt{n}$. Then we have

$$\frac{1}{\log n} \sum_{i=1}^{n} \frac{1}{i} \delta\{\hat{s}_i(\cdot)\} \Longrightarrow W \quad a.s. \tag{7}$$

with uniform partition. Here $\delta(x)$ is the point mass at $x \in D[0,1]$ and W is the standard Wiener process.

The statement (7) for i.i.d random variables with mean 0 was proved under the condition of finite $(2+\varepsilon)$th moments by Schatte [7] for $\varepsilon = 1$, by Brosamler [3] for $\varepsilon > 0$ and by Lacey and Philipp [4] for $\varepsilon = 0$.

Peligrad and Shao [6] proved

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} I_A\left(\frac{S_k}{\sigma_k}\right) = \frac{1}{\sqrt{2\pi}} \int_A e^{-\frac{u^2}{2}}\, du \quad \text{a.s.} \tag{7*}$$

where $\sigma_n^2 = \mathbf{E}S_n^2$, for weakly dependent random variables under certain conditions, including our conditions in the Corollary. The statement (7) is stronger than (7*).

2. Proofs

In the proof we shall apply the following two lemmas.

Lemma 1 (Corollary 2.7, Utev [8]). *Let* $1 \leqq t \leqq 2$ *and* $\phi = \sum_{k=1}^{\infty} \phi^{\frac{1}{2}}(2^k) < \infty$; $\mathbf{E}X_i = 0, \quad i = 1, 2, \ldots, n$. *Then*

$$\mathbf{E} \max_{1 \leqq i \leqq n} |S_i|^t \leqq k(\phi) \sum_{i=1}^{n} \mathbf{E}|X_i|^t,$$

where $k(\phi)$ *depends only on* ϕ.

Lemma 2. *Let* $\xi_1, \xi_2, \ldots$ *be an arbitrary sequence of real valued random variables such that* $\mathbf{E}\xi_i = 0$ *and* $\mathbf{E}\xi_i^2 < C$ *for all* $i \geqq 1$. *If*

$$\mathbf{E}\left(\sum_{i=1}^{n} \frac{1}{i}\xi_i\right)^2 = O(\log n) \tag{8}$$

then we have

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{i=1}^{n} \frac{1}{i}\xi_i = 0 \quad a.s.. \tag{9}$$

The proof of Lemma 2 can be given by applying Chebyshev's inequality for the subsequence $n_k = e^{k^2}$ (see Peligrad and Shao [6] or Schatte [7]).

Proof of Theorem. Let $BL(D[0,1])$ denote the set of the functions $g : D[0,1] \longrightarrow \mathbf{R}$ such that for some $C > 0$

$$|g(x) - g(y)| \leqq C d(x,y), \quad |g(x)| \leqq C \qquad \text{for all } x, y \in D[0,1] \tag{10}$$

where d is the Skorohod metric. To prove Theorem 1 it is enough to show that for any $g \in BL(D[0,1])$

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{i=1}^{n} \frac{1}{i}\xi_i = 0 \quad \text{a.s.}$$

where $\xi_i = g(\hat{s}_i(\cdot)) - \mathbf{E}g(\hat{s}_i(\cdot))$ (see [2]). In view of Lemma 2, if we show that (8) holds for $\{\xi_k, k \geqq 1\}$ then our theorem will be proved. For this purpose we define for any $k < l$, the function $s^*_{k,l} : [0,1] \longrightarrow \mathbf{R}$ by

$$s^*_{k,l}(t) = \begin{cases} 0, & \text{if} \quad 0 \leqq t \leqq t_{l,k} \\ \dfrac{S_j - S_k}{a_l}, & \text{if} \quad t_{l,j} \leqq t \leqq t_{l,j+1} \quad (k \leqq j \leqq l-1), \end{cases}$$

where a_n is as in (1). It is clear that $d(\hat{s}_l, s^*_{k,l}) \leqq \|\hat{s}_l - s^*_{k,l}\|_\infty$ and thus by (10)

$$|g(\hat{s}_l) - g(s^*_{2k,l})| \leqq C \frac{\max\limits_{1 \leqq j \leqq 2k} |S_j|}{a_l} \quad \text{for } 2k < l. \tag{11}$$

Now we consider the covariance of ξ_k and ξ_l for $2k < l$:

$$\begin{aligned} |\mathbf{E}(\xi_k \xi_l)| &= |\mathbf{Cov}(g(\hat{s}_k), g(\hat{s}_l))| \leqq \\ &\leqq |\mathbf{Cov}(g(\hat{s}_k), g(\hat{s}_l) - g(s^*_{2k,l}))| + |\mathbf{Cov}(g(\hat{s}_k), g(s^*_{2k,l}))|. \end{aligned}$$

Here

$$|\mathbf{Cov}(g(\hat{s}_k), g(s^*_{2k,l}))| \leqq C \phi^{\frac{1}{2}}(k) \tag{12}$$

since $g(\hat{s}_k) \in M_1^k$, $g(s^*_{2k,l}) \in M_{2k}^\infty$ (see Billingsley [1], p. 171, Lemma 2). Further by (10) and (11) we can see that

$$|\mathbf{Cov}(g(\hat{s}_k), g(\hat{s}_l) - g(s^*_{2k,l}))| \leqq C \frac{\mathbf{E}\left(\max\limits_{1 \leqq j \leqq 2k} |S_j|\right)}{a_l}. \tag{13}$$

The condition (3), Lemma 1 and Cauchy–Schwarz inequality imply that

$$\mathbf{E}(\max_{1 \leqq j \leqq 2k} |S_j|) \leqq (\mathbf{E} \max_{1 \leqq j \leqq 2k} |S_j|^2)^{\frac{1}{2}} \leqq C a_{2k}. \tag{14}$$

Hence from (12), (13) and (14) it follows that

$$|\mathbf{E}(\xi_k \xi_l)| \leqq C_1 \frac{a_{2k}}{a_l} + C_2 \phi^{\frac{1}{2}}(k) \quad \text{for } 2k < l.$$

Notice that

$$\begin{aligned} &\mathbf{E}\left(\sum_{i=1}^n \frac{1}{i} \xi_i\right)^2 = \\ &= \sum_{k=1}^n \frac{1}{k^2} \mathbf{E}|\xi_k|^2 + 2 \sum_{\substack{1 \leqq k < l \leqq n \\ 2k \geqq l}} \frac{|\mathbf{E}(\xi_k \xi_l)|}{kl} + 2 \sum_{\substack{1 \leqq k < l \leqq n \\ 2k < l}} \cdot \frac{|\mathbf{E}(\xi_k \xi_l)|}{kl} = \\ &= \sum\nolimits_1 + \sum\nolimits_2 + \sum\nolimits_3. \end{aligned} \tag{15}$$

Now by easy calculation we can see that

$$\sum\nolimits_1 \leqq \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty, \tag{16}$$

$$\sum\nolimits_2 \leqq C \sum_{k=1}^{n} \sum_{l=k+1}^{2k} \frac{1}{kl} = O(\log n), \tag{17}$$

$$\begin{aligned} \sum\nolimits_3 \leqq C_2 \sum_{l=2}^{n} \frac{1}{l} \sum_{k=1}^{l-1} \frac{\phi^{\frac{1}{2}}(k)}{k} &+ \\ + C_1 \sum_{l=2}^{n} \sum_{k=1}^{l-1} \frac{1}{k^{1-\gamma} l^{1+\gamma}} &= O(\log n), \end{aligned} \tag{18}$$

on account of (2), (4), (12) and boundedness of ξ_i. Hence (16)–(18) together imply (8). This completes the proof of the Theorem.

Proof of Corollary. It follows from Corollary 2.4 of [5] that under the conditions of our Corollary the weak convergence

$$\frac{S_{[nt]}}{\sigma\sqrt{n}} \Longrightarrow W(t)$$

holds which in turn implies (6) with $a_n = \sigma\sqrt{n}$, uniform partition and the Wiener measure μ. Thus our Corollary immediately follows from Theorem 1.

Acknowledgements. I am indebted to E. Csáki and to the referee for their helpful remarks.

REFERENCES

[1] Billingsley, P., *Convergence of probability measures*, Wiley, New York, 1968. *MR* **38** #1718

[2] Berkes, I. and Dehling, H., Some limit theorems in log density, *Ann. Probab.* **21** (1993), 1640–1670. *MR* **94h**:60026

[3] Brosamler, G. A., An almost everywhere central limit theorem, *Math. Proc. Cambridge Philos. Soc.* **104** (1988), 561–574. *MR* **89i**:60045

[4] Lacey, M. T. and Philipp, W., A note on the almost sure central limit theorem, *Statist. Probab. Lett.* **9** (1990), 201–205. *MR* **91e**:60100

[5] Peligrad, M., Invariance principles for mixing sequences of random variables, *Ann. Probab.* **10** (1982), 968–981. *MR* **84c**:60054

[6] Peligrad, M. and Shao, Q. M., A note on the almost sure central limit theorem for weakly dependent random variables, *Statist. Probab. Lett.* **22** (1995), 131–136.

[7] Schatte, P., On strong versions of the central limit theorem, *Math. Nachr.* **137** (1988), 249–256. *MR* **89i**:60070

[8] UTEV, S. A., Sums of ϕ-mixing random variables, *Trudy Inst. Mat. (Novosibirsk)* **13** (1989), Asimptot Analiz Raspred. Sluch. Protsess, 78–100 (in Russian). *MR* **91**g:60043

(Received February 20, 1993)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 203–224

LIMIT THEOREMS ABOUT THE DISTRIBUTION OF ALMOST PERIODIC FUNCTIONS

P. MAJOR

Dedicated to Paul Révész for his sixtieth birthday

Abstract

We prove a limit theorem about the distribution of an almost periodic function $F(R) = \sum_{n=1}^{\infty} a_n e^{2\pi i \lambda_n R}$, $\sum_{n=1}^{\infty} |a_n|^2 < \infty$, when R is uniformly distributed in an interval $[0,T]$, and $T \to \infty$. Also a limit theorem is proved about the distribution of the random vector $(F(R), F(R+w(R,T)))$, $R \in [0,T]$, if the function $w(R,T)$ is appropriately defined. Similar results were proved also in other papers. (See [2] and [3].) The proofs in this paper are essentially different from the previous ones, and they may give some new insight to this problem. Previous proofs were based on ergod theoretical arguments, while in this paper some standard methods of Fourier analysis are applied. These investigations were motivated by the study of the limit behaviour of the number of lattice points in a randomly magnified strip in the plane.

1. Introduction

In this paper the following problem is discussed: Let us consider a function

$$F(R) = \sum_{n=1}^{\infty} a_n e^{2\pi i \lambda_n R} \tag{1.1}$$

with

$$\sum_{n=1}^{\infty} |a_n|^2 < \infty\,, \tag{1.2}$$

where $\lambda_1, \lambda_2, \dots$ are different non-zero real numbers. We also assume that $\lambda_{2n} = -\lambda_{2n-1}$ and $a_{2n} = \bar{a}_{2n-1}$ for $n = 1, 2, \dots$. This restriction is not essential, we only impose it to work with real valued functions. Define the distribution μ_T of the function $F(R)$ with respect to the uniform distribution in the interval $[0,T]$ by the formula

$$\mu_T(\mathbf{A}) = \frac{1}{T}\lambda\{R\colon 0 \leqq R \leqq T,\ F(R) \in \mathbf{A}\} \tag{1.3}$$

1991 *Mathematics Subject Classification.* Primary 11K70; Secondary 60F05, 11P21.

Key words and phrases. Limit theorems, Besicovich spaces, weak convergence in probability theory, Fourier analysis.

for any measurable set $\mathbf{A} \subset \mathbb{R}^1$, where λ denotes the Lebesgue measure. We want to prove that the measures μ_T have a weak limit. We also want to prove a generalization of this result in the case when the joint distribution of the functions $F(R)$ and $F(R+w(R,T))$ are investigated with a nice function $w(R,T)$. We shall study the limit distribution of this vector if R is uniformly distributed in an interval $[aT, bT]$, $0<a<b\leqq 1$, and $T\to\infty$. Choose some constants $0<a<b\leqq 1$, consider a function $w(R,T)$, $aT\leqq R\leqq bT$, and define the joint distribution of the functions $F(R)$ and $F(R+w(R,T))$ by the formula

(1.4)
$$\mu_{T,w,(a,b)}(\mathbf{A}) = \frac{1}{(b-a)T}\lambda\{R\colon aT\leqq R\leqq bT,\ (F(R), F(R+w(R,T)))\in \mathbf{A}\},$$

for any measurable set $\mathbf{A}\subset\mathbb{R}^2$. We want to prove that under appropriate conditions the measures $\mu_{T,w,(a,b)}$ with fixed numbers $0<a<b\leqq 1$ converge weakly to a probability measure as $T\to\infty$.

Problems of such type arose in the investigation of the number of lattice points in a randomly magnified domain $R\mathbf{C}$, where $\mathbf{C}$ is a convex set with a smooth boundary, and R is a randomly chosen magnifying constant. It is proved, (see [2]), that the number of lattice points $N(R)$ in the domain $R\mathbf{C}$ after an appropriate normalization $\chi(R) = \dfrac{N(R)-\text{Area}(R\mathbf{C})}{\sqrt{R}}$ can be written in the form (1.1). We are interested in the limit behaviour of the number of lattice points in a randomly enlarged domain $R\mathbf{C}$ or in a randomly defined strip $(R+\alpha(R))\mathbf{C}\setminus R\mathbf{C}$, with an appropriately defined function $\alpha(R)$, when the number R is randomly chosen. This can be described by means of the representation of $\chi(\mathbf{C})$ in the form of a series (1.1) and the above indicated limit theorems.

Actually the results of the present paper are only slight generalizations of earlier papers (see [2], [3]), where similar results were proved because of the same motivation. Nevertheless, we think that it is useful to revisit this problem for the following reason: Our approach is different from that of the above mentioned papers, and we think that it has some interesting features. In previous papers the proofs were based on the ergod theorem. Because of this approach some measure theoretical problems arose whose solution seems to be hard. We want to show that these problems can be avoided by replacing the ergod theorem by a multi-dimensional continuous time version of the following well-known number theoretical result: For an irrational number α the sequence $n\alpha \pmod 1$, $n=1,2,\dots$, is asymptotically uniformly distributed in the interval $[0,1]$.

Before formulating the results of this paper we have to explain the content of formula (1.1). The function $F(R)$ in this formula is considered as an element of the Besicovitch space, i.e. we assume that it has the following property: For all $\varepsilon > 0$ there exists an index $p_0 = p_0(\varepsilon)$ in such a way that

$$\limsup_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \left| F(R) - \sum_{n=1}^{p} a_n e^{2\pi i \lambda_n R} \right|^2 dR < \varepsilon \tag{1.5}$$

for $p > p_0$. The theory of Besicovitch spaces can be found in [1], but in the present paper we do not need its fine details. Here we only use relation (1.5). Let us remark that the definition in (1.5) does not define the function $F(R)$ in a unique way. Indeed, if $F(R)$ and $\bar{F}(R)$ are two functions such that

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} |F(R) - \bar{F}(R)|^2 \, dR = 0 \,,$$

then the functions $F(R)$ and $\bar{F}(R)$ simultaneously satisfy or do not satisfy relation (1.5). Thus the theorems formulated below state in particular that the limit distribution appearing in them do not depend on which function $F(R)$ we take from those satisfying formula (1.5). Our first result is the following

THEOREM 1. *For all functions $F(R)$ satisfying* (1.1) *and* (1.2) *the probability measures μ_T defined in* (1.3) *converge weakly to a probability measure μ as $T \to \infty$. Moreover, for all continuous functions $g(u)$ such that $|g(u)| < Au^2 + B$ with some appropriate numbers $A > 0$ and $B > 0$, the relation*

$$\lim_{T\to\infty} \frac{1}{T} \int_{0}^{T} g(F(R)) \, dR = \lim_{T\to\infty} \int g(u)\mu_T(du) = \int g(u)\mu(du) \tag{1.6}$$

holds. In particular,

$$\lim_{T\to\infty} \frac{1}{T} \int_{0}^{T} F(R) \, dR = 0 \,, \tag{1.7}$$

and

$$\lim_{T\to\infty} \frac{1}{T} \int_{0}^{T} F(R)^2 \, dR = \sum_{n=1}^{\infty} |a_n|^2 \,. \tag{1.7'}$$

To formulate Theorem 2 first we have to clarify how to define the class of "width" functions $w(R,T)$ in it. We consider two different cases. In case (a) this width has constant order, and in a point $R = uT$, $a \leqq u \leqq b$, the value of the "width" function $w(R,T)$ is close to a monotone function $K(u)$ for large T, and in case (b) it tends to infinity as $T \to \infty$ in a regular way. First we formulate Theorem 2, and then show that it contains the results of Section 2 in [3] as a special case.

THEOREM 2. *Let a function $F(R)$ be given, which satisfies relations* (1.1) *and* (1.2), *and let $K(x)$ be a continuously differentiable, monotone (increasing or decreasing) function with non-vanishing derivative in an interval $[a,b]$, $0<a<b\leqq 1$, and $K(x)>0$ for all $x\in[a,b]$. Assume that the function $w(R,T)$, $T>0$, $aT<R<bT$, satisfies one of the following conditions:*

(a) $w(R,T)=K\left(\frac{R}{T}\right)+o(1)$ *and*

$$\frac{\partial}{\partial R}w(R,T)=\frac{1}{T}K'\left(\frac{R}{T}\right)(1+o(1)),\quad aT\leqq R\leqq bT.$$

(b) *There exists some function $L(T)$, $L(T)\to\infty$ and $\frac{L(T)}{T}\to 0$ as $T\to\infty$, such that $w(R,T)=L(T)K\left(\frac{R}{T}\right)(1+o(1))$, and*

$$\frac{\partial}{\partial R}w(R,T)=\frac{L(T)}{T}K'\left(\frac{R}{T}\right)(1+o(1)),\quad aT\leqq R\leqq bT.$$

The term $o(1)$ is uniformly small for $aT\leqq R\leqq bT$ as $T\to\infty$ in both cases (a) *and* (b).

Then the measures $\mu_{T,w,(a,b)}$ defined in (1.4) *with these functions $w(R,T)$ have a weak limit $\bar\mu$ on $\mathbb{R}^2$ as $T\to\infty$. In case* (a) *$\bar\mu$ equals some probability measure $\mu_{(a,b)}^{K(x)}$, i.e. it depends only on the function $K(x)$ and not on the special form of the function $w(R,T)$. The relation*

$$\bar\mu=\mu_{(a,b)}^{\infty}=\mu\times\mu \tag{1.8}$$

holds, if $w(R,T)$ satisfies the conditions of case (b), *where μ is the probability measure defined in Theorem 1, and $\times$ denotes direct product. In particular, the limit measures $\mu_{(a,b)}^{\infty}$ do not depend on the parameters a and b.*

The statement about the weak convergence of the measures $\mu_{T,w,(a,b)}$ can be strengthened in the following way: If $w(R,T)$ satisfies condition (a) *or* (b), *and $g(u,v)$ is a continuous function such that $|g(u,v)|<A(u^2+v^2)+B$ with some appropriate $A>0$ and $B>0$, then*

$$\lim_{T\to\infty}\frac{1}{(b-a)T}\int_{aT}^{bT}g(F(R),F(R+w(R,T))\,dR=$$
$$\lim_{T\to\infty}\int g(u,v)\,\mu_{T,w,(a,b)}(du,dv)=\int g(u,v)\,\bar\mu(du,dv) \tag{1.9}$$

with $\bar\mu=\mu_{(a,b)}^{K(x)}$ in case (a) *and $\bar\mu=\mu_{(a,b)}^{\infty}=\mu\times\mu$ in case* (b).

Let us fix some function $K(x)$ which satisfies the conditions imposed on it in Theorem 2. Then, the probability measures $\mu_{(a,b)}^{zK(x)}$ depend continuously on z in the weak topology for $0 < z < \infty$, and

$$\lim_{z\to\infty} \mu_{(a,b)}^{zK(x)} = \mu_{(a,b)}^{\infty}. \tag{1.10}$$

Let us recall that the probability measures $\mu^{zK(x)}$ are called continuous in the weak topology if for all bounded and continuous functions g the integrals $\int g\, d\mu^{zK(x)}$ are continuous functions of z, and they converge to a measure μ^{∞} as $z \to \infty$ if $\lim_{z\to\infty} \int g\, d\mu^{zK(x)} = \int g\, d\mu^{\infty}$ for all bounded and continuous functions g.

In paper [3] the following problem is investigated. Let the function $F(R)$ be equal to the normalized number of lattice points $\chi(R)$ in a domain $R\mathbf{C}$. We are interested in the asymptotic behaviour of the number of lattice points in a strip $(R+w(R,T))\mathbf{C}\setminus R\mathbf{C}$, where R is uniformly distributed in an interval $aT < R < bT$. In such an investigation the knowledge of the limit distribution of the vector $(F(R), F(R+w(R,T)))$, $aT \leqq R \leqq bT$, as $T \to \infty$, can be useful. The most interesting choice of the function $w(R,T)$ is that when for fixed T the area of the set $(R+w(R,T))\mathbf{C}\setminus R\mathbf{C}$ equals a constant $S(T)$, and the function $S(T)$ satisfies the relation

$$\lim_{T\to\infty} \frac{S(T)}{2T} = z, \quad 0 < z \leqq \infty, \qquad \text{and} \quad \lim_{T\to\infty} \frac{S(T)}{T^2} = 0. \tag{1.11}$$

This case is considered in paper [3]. If the area of the set $\mathbf{C}$ equals one, and the area of the strip $(R+w(R,T))\mathbf{C}\setminus R\mathbf{C}$ is $S(T)$, then the function $w(R,T)$ satisfies the equality

$$w(R,T)^2 + 2Rw(R,T) = S(T). \tag{1.12}$$

It is not difficult to see that the function $w(R,T)$ defined by formulas (1.11) and (1.12) satisfies the conditions of Theorem 2. If the number z is finite, then case (a) of Theorem 2 holds with $K(x) = z/x$, and if it equals infinity, then case (b) holds with $K(x) = 1/x$ and $L(T) = S(T)/2T$. Hence the results of Section 2 of [3] are consequences of Theorem 2 with the above choice of the function $K(x)$ and $L(T)$.

One would like to give an explicit description of the limit measures appearing in Theorems 1 and 2. We return to this question at the end of this paper. Here we formulate a result which gives a decomposition of the measures $\mu_{(a,b)}^{K(x)}$. Let us define the distribution of the vector $(F(R), F(R+x))$ in the interval $[0,T]$ with a fixed number $0 \leqq x < \infty$ by the formula:

$$\nu_T^x(\mathbf{A}) = \frac{1}{T}\lambda\{R\colon 0 \leqq R \leqq T,\ (F(R), F(R+x)) \in \mathbf{A}\} \tag{1.13}$$

for any measurable set $\mathbf{A} \subset \mathbb{R}^2$. Now we formulate the following

THEOREM 3. *For fixed $0<x<\infty$ the measures ν_T^x converge weakly to a probability measure ν^x, and also the relation*

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T g(F(R),F(R+x))\,dR=\int g(u,v)\,\nu^x(du,dv)\,, \tag{1.14}$$

holds if $g(u,v)$ is a continuous function, and $|g(u,v)|<A(u^2+v^2)+B$ with some constants $A>0$ and $B>0$. For a fixed function $g(u,v)$ the integral at the right-hand side of (1.14) *is a continuous and bounded function of x.*

The identity

$$\mu_{(a,b)}^{K(x)}=\frac{1}{(b-a)}\int_a^b \nu^{K(x)}\,dx=\frac{1}{(b-a)}\int_{K(a)}^{K(b)}\frac{\nu^x}{K'(K^{-1}(x))}\,dx \tag{1.15}$$

holds for the function $\mu_{a,b}^{K(x)}$ defined in Theorem 2.

We shall prove the following corollary of the above results:

COROLLARY. *Let $h(x)$ be an integrable function on an interval $[a,b]$, $0<$ $<a<b\leqq 1$. Let the function $w(R,T)$ satisfy the conditions of case (a) of Theorem 2. Then the relation*

$$\begin{aligned}&\lim_{T\to\infty}\frac{1}{T}\int_{aT}^{bT} g(F(R),F(R+w(R,T)))h\left(\frac{R}{T}\right)dR=\\ &\qquad=\int_a^b h(x)\int g(u,v)\,\nu^{K(x)}(du,dv)\,dx\end{aligned} \tag{1.16}$$

holds for all continuous functions $g(u,v)$ if one of the following conditions is satisfied: Either $g(u,v)$ is bounded or $h(u)$ is square integrable and $|g(u,v)|<$ $<A(u^2+v^2)+B$ with some appropriate numbers $A>0$ and $B>0$.

This paper consists of four sections. In Section 2 we prove the Theorems in the special case when the sum (1.1) defining the function $F(R)$ contains finitely many terms. In Section 3 we carry out a limiting procedure which proves the results of the paper by means of Section 2. In Section 4 we make some comments and prove some generalizations.

Let us make a short comparison of the method of this paper with previous ones. The main difference of the proof of Theorem 1 in this paper and in [2] is that we replace the application of the ergod theorem by a number theoretical distribution theorem. The formulation of Theorems 2 and 3 are very close to the results of Section 2 in [3]. The proofs are essentially different. In paper [3] these results were deduced from Theorem 1 by a tricky ergod theoretical argument. Here we show that a slight modification of the proof of Theorem 1 supplies a direct proof for them.

2. The proof of the results in a special case

Let us consider the case when the function $F(R)$ is defined by a finite sum

$$F(R) = F_p(R) = \sum_{n=1}^{p} a_n e^{2\pi i \lambda_n R} \tag{2.1}$$

with an even number p, real non-zero numbers λ_n such that $\lambda_{2n} = -\lambda_{2n-1}$, $a_{2n} = \bar{a}_{2n-1}$, and the measures μ_T, $\mu_{T,w,(a,b)}$ and ν_T^x are defined in formulas (1.3), (1.4) and (1.13) by means of this function. We prove in this Section Theorems 1, 2 and 3 in the case when these measures are determined by a function of the form (2.1). In the next Section we prove the result for general functions $F(R)$ defined in (1.1) by approximating them with the functions F_p appearing in (2.1). We shall indicate the dependence of these measures on the functions F_p by denoting them as $\mu_T(F_p)$, $\mu_{T,w,(a,b)}(F_p)$ and $\nu_T^x(F_p)$ when necessary. The main results of this Section are the following

PROPOSITIONS 1, 2 AND 3. *Let the function $F(R)$ be defined by the finite trigonometrical sum* (2.1), *and let the measures μ_T, $\mu_{T,w,(a,b)}$ and ν_T^x be defined in formulas* (1.3), (1.4), *and* (1.13) *by means of this function $F(R)$. If the function $w(R,T)$ and $S(T)$ satisfies the conditions of Theorem 2, then Theorems 1, 2, and 3 hold with this choice of the corresponding measures.*

PROOF OF PROPOSITION 1. We have to investigate the asymptotic behaviour of the expression

$$\frac{1}{T}\int_0^T g(F(R))\,dR \tag{2.2}$$

as $T \to \infty$ in the case when $g(u)$ is a bounded continuous function. We shall rewrite, following the argument of [2] and [4], the expression in (2.2) as an integral on a torus with respect to an appropriate measure. It is useful to work, when handling the function $F(R)$, with frequencies linearly independent over the rational numbers. Since the frequencies λ_n may not have this property we express them as a linear combination of some numbers $\tau_1, \dots, \tau_s$ linearly independent over the rational numbers

$$\lambda_n = T_n(\tau_1, \dots, \tau_s) = \sum_{k=1}^{s} A(n,k)\tau_k\,, \quad n = 1, 2, \dots, p,\ k = 1, \dots, s \tag{2.3}$$

with integer coefficients $A(n,k)$. Let V denote the unit interval with the group action addition modulo 1. Introduce its s-fold and p-fold direct products

$$\mathcal{V} = \underbrace{V \times \cdots \times V}_{s\ \text{times}} \tag{2.4}$$

and

$$\mathcal{V}' = \underbrace{V \times \cdots \times V}_{p \text{ times}} . \tag{2.4'}$$

Define the maps $U \colon \mathbb{R}^1 \to \mathcal{V}$

$$U(R) = \{R\tau_k \ (\text{mod}\, 1), \ k = 1, \ldots, s\} . \tag{2.5}$$

$V \colon \mathcal{V} \to \mathcal{V}'$

$$V(u_1, \ldots, u_s) = \left\{ \sum_{n=1}^{s} A(n,k) u_k \ (\text{mod}\, 1), \ n = 1, \ldots, p \right\} \tag{2.6}$$

for $(u_1, \ldots, u_s) \in \mathcal{V}$ with the integer coefficients $A(n,k)$ appearing in (2.3) and $G \colon \mathcal{V}' \to \mathbb{R}^1$

$$G(u_1, \ldots, u_p) = \sum_{n=1}^{p} a_n e^{2\pi i u_n} . \tag{2.7}$$

Clearly, $F(R) = G(V(U(R)))$. Define the probability measure ρ_T on $\mathcal{V}$ induced by the map U by the formula

$$\rho_T(\mathbf{A}) = \frac{1}{T} \lambda\{R \colon 0 \leqq R \leqq T, \ U(R) \in \mathbf{A}\} \tag{2.8}$$

for all measurable sets $\mathbf{A} \subset \mathcal{V}$.

Then the integral (2.2) can be rewritten as

$$\frac{1}{T} \int_0^T g(F(R))\, dR = \int_{\mathcal{V}} g(G(V(u)))\, \rho_T(du) . \tag{2.9}$$

The relation

$$\rho_T \Rightarrow \rho \quad \text{as } T \to \infty \tag{2.10}$$

holds, where ρ denotes the Haar measure on $\mathcal{V}$, and $\Rightarrow$ means weak convergence of probability measures. Relation (2.10) is a known result. Nevertheless, we give its proof, because it is short, and we need its modification in the proof of Proposition 2. By Weil's lemma (or by the characteristic function method on commutative compact groups) to prove (2.10) it is enough to check that, with the notation $(u_1, \ldots, u_s) = u \in \mathcal{V}$,

$$\lim_{T\to\infty} \int \exp\left\{ 2\pi i \sum_{k=1}^{s} m_k u_k \right\} \rho_T(du) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \exp\left\{ 2\pi i \sum_{k=1}^{s} m_k \tau_k R \right\} dR$$

$$= \lim_{T\to\infty} \frac{\exp\left\{ 2\pi i T \sum\limits_{k=1}^{s} m_k \tau_k \right\} - 1}{2\pi i T \sum\limits_{k=1}^{s} m_k \tau_k} = 0$$

if $m_1, m_2, \dots, m_k$ are integers, and not all of them equal zero. Relations (2.9) and (2.10) imply that

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T g(F(R))\, dR = \lim_{T\to\infty} \int g(G(V(u))\rho_T(du) = \int g(G(V(u))\rho(du)\,,$$

since $g(G(V(u)))$ is a bounded, continuous function. The last relation implies that

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T g(F(R))\, dR = \int g(u)\,\mu(du) \tag{2.11}$$

with the measure μ defined on $\mathbb{R}^1$ by the relation

$$\mu(\mathbf{A}) = \rho\{u\colon u \in \mathcal{V},\ G(V(u)) \in \mathbf{A}\} \tag{2.12}$$

for all measurable sets $\mathbf{A} \in \mathbb{R}^1$. Relations (2.11) and (2.12) imply that the measures μ_T converge weakly to the measure μ defined in (2.12). To complete the proof of Proposition 1 observe that the function $|F(R)|$ is bounded by $C_p = \sum_{n=1}^{p} |a_n|$ for all $R \in \mathbb{R}^1$. Hence all measures μ_T and μ are concentrated in the interval $[-C_p, C_p]$, and relation (1.6) follows for all continuous functions $g(u)$, since they can be replaced by their truncation at $\pm C_p$, which are bounded continuous functions. Finally, relations (1.7) and (1.7′) follow from the observation that $F(R)$ is a finite sum, and the relations

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T e^{i\lambda_n R}\, dR = 0$$

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T e^{i(\lambda_n - \lambda_{n'})R}\, dR = \delta(n, n')$$

hold. □

Proof of Proposition 2. The proof of Proposition 2 is very similar to that of Proposition 1. The main difference is that we have to carry out the integral transformations by means of different functions G, U and V, and the statement formulated in (2.10) has to be generalized.

Define the maps $U\colon \mathbb{R}^1 \to \mathcal{V} \times \mathcal{V}$

$$\begin{aligned} U(R) = U(R, w, T) = \{&R\tau_k \ (\mathrm{mod}\ 1),\ k = 1, \dots, s\,, \\ &(R + w(R, T))\tau_k \ (\mathrm{mod}\ 1),\ k = 1, \dots, s\}\,, \end{aligned} \tag{2.13}$$

$V: \mathcal{V} \times \mathcal{V} \to \mathcal{V}' \times \mathcal{V}'$

$$V(u_1, \dots, u_{2s}) = \left\{ \sum_{k=1}^{s} A(n,k) u_k \ (\mathrm{mod}\, 1), \quad n = 1, \dots, p, \right. \tag{2.14}$$
$$\left. \sum_{k=1}^{s} A(n,k) u_{k+s} \ (\mathrm{mod}\, 1), \quad n = 1, \dots, p \right\}$$

for $(u_1, \dots, u_{2s}) \in \mathcal{V} \times \mathcal{V}$ with the integer coefficients $A(n,k)$ appearing in (2.3) and $G: \mathcal{V}' \times \mathcal{V}' \to \mathbb{R}^2$

$$G(u_1, \dots, u_{2p}) = \left(\sum_{n=1}^{p} a_n e^{2\pi i u_n}, \ \sum_{n=1}^{p} a_n e^{2\pi i u_{n+p}} \right). \tag{2.15}$$

Then $(F(R), F(R + w(R,T))) = G(V(U(R)))$. Define the probability measure $\bar\rho_T = \bar\rho_{T,w,(a,b)}$ on $\mathcal{V} \times \mathcal{V}$ induced by the map U by the formula

$$\bar\rho_{T,w,(a,b)}(\mathbf{A}) = \frac{1}{(b-a)T} \lambda\{R: aT \leqq R \leqq bT, \ U(R,w,T)) \in \mathbf{A}\} \tag{2.16}$$

for any measurable set $\mathbf{A} \subset \mathcal{V} \times \mathcal{V}$.

We claim that if $w(R,T)$ satisfies the conditions of Theorem 2, then the limit relation

$$\bar\rho_{T,w,(a,b)} \Rightarrow \bar\rho \quad \text{as } T \to \infty \tag{2.17}$$

holds with a probability measure $\bar\rho$ on $\mathcal{V} \times \mathcal{V}$. Moreover, we claim that

$$\bar\rho = \bar\rho_{(a,b)}^{K(x)} \tag{2.17'}$$

if $w(R,T)$ satisfies the conditions of case (a) of Theorem 2 with $K(x)$, i.e. the limit depends only on the function $K(x)$ in this case, and

$$\bar\rho = \bar\rho_{a,b}^{\infty} = \text{the Haar measure } \rho \times \rho \text{ on } \mathcal{V} \times \mathcal{V} \tag{2.17''}$$

if $w(R,T)$ satisfies the conditions of case (b) of Theorem 2.

To prove (2.17) we show that the Fourier coefficients

$$L_{w,T}(m_1, \dots, m_{2s}) = \int_{\mathcal{V} \times \mathcal{V}} \exp\left\{ 2\pi i \sum_{k=1}^{2s} m_k u_k \right\} \bar\rho_{T,w,(a,b)}(du)$$

with $u = (u_1, \dots, u_{2s}) \in \mathcal{V} \times \mathcal{V}$ have a limit

$$\lim_{T \to \infty} L_{w,T}(m_1, \dots, m_{2s}) = L(m_1, \dots, m_{2s}) \tag{2.18}$$

for all integers $m_1, \ldots, m_{2s}$. These Fourier coefficients can be rewritten by an integral transformation as
(2.19)

$$L_{w,T}(m_1, \ldots, m_{2s}) = \frac{1}{(b-a)T} \int_{aT}^{bT} e^{i(A(m_1,\ldots,m_{2s})R + B(m_1,\ldots,m_{2s})w(R,T))}\, dR$$

with

$$(2.19') \qquad \begin{aligned} A(m_1, \ldots, m_{2s}) &= 2\pi \sum_{k=1}^{s} (m_k + m_{k+s})\tau_k\,, \\ B(m_1, \ldots, m_{2s}) &= 2\pi \sum_{k=1}^{s} m_{k+s}\tau_k\,. \end{aligned}$$

Because of the linear independence of the numbers τ_k both expressions $A(m_1, \ldots, m_{2s})$ $B(m_1, \ldots, m_{2s})$ can disappear simultaneously only if all coefficients m_k are zero, which is a trivial case. Otherwise we claim that

$$(2.20) \qquad \begin{aligned} &\lim_{T\to\infty} L_{w,T}(m_1, \ldots, m_{2s}) \\ &= \begin{cases} 0 & \text{if } A(m_1, \ldots, m_{2s}) \neq 0 \\ \frac{1}{(b-a)} \int_a^b e^{iBK(u)}\, du & \text{if } A(m_1, \ldots, m_{2s}) = 0 \text{ and} \\ & \quad w(R,T) \text{ satisfies case (a)} \\ 0 & \text{if } A(m_1, \ldots, m_{2s}) = 0 \text{ and} \\ & \quad w(R,T) \text{ satisfies case (b).} \end{cases} \end{aligned}$$

The first line in relation (2.20) can be proved by means of relation (2.19) with the change of variables $AR + Bw(R,T) = u$. Let us observe that because of the conditions of Theorem 2 $\frac{du}{dR} = A + o(1)$ uniformly for $aT \leq R \leq bT$, and the boundaries of the domain of integration after the change of variables are $aAT(1+o(1))$ and $bAT(1+o(1))$. Hence we get that

$$\begin{aligned} &\lim_{T\to\infty} L_{w,T}(m_1, \ldots, m_{2s}) \\ &= \lim_{T\to\infty} \frac{(1+o(1))}{(b-a)A(m_1, \ldots, m_{2s})T} \int_{aA(m_1,\ldots,m_{2s})T}^{bA(m_1,\ldots,m_{2s})T} e^{iu}\, du = 0 \end{aligned}$$

in this case. If the conditions of the second line of (2.20) hold, i.e. when the conditions of case (a) of Theorem 2 hold, and $A = 0$, then we can calculate the expression (2.19) with the change of variables $u = \frac{R}{T}$. Simple calculation shows that

$$\frac{1}{(b-a)T}\int_{aT}^{bT} e^{iBw(R,T)}\,dR \to \frac{1}{(b-a)}\int_a^b e^{iBK(u)}\,du\,,$$

and the second relation of (2.20) holds. The third line of (2.20) (this case holds when $A=0$ and condition (b) is satisfied) can be proved similarly with the change of variables $u=\frac{R}{T}$ and $v=\frac{1}{L(T)}w(uT,T)$. Some calculation shows that $v=K(u)(1+o(1))$, $\frac{\partial v}{\partial u}=K'(u)(1+o(1))$, and

$$\begin{aligned}\frac{1}{(b-a)T}\int_{aT}^{bT} e^{iBw(R,T)}\,dR &= \frac{1}{(b-a)}\int_a^b e^{iBw(uT,T)}\,du\\ &= \frac{1}{(b-a)}\int_{K(a)}^{K(b)} e^{iBL(T)v}\frac{1}{K'(K^{-1}(v))}(1+o(1))\,dv \to 0\end{aligned}$$

by the Riemann lemma. The convergence of the Fourier coefficients formulated in relation (2.20) implies formulas (2.17), (2.17′) and (2.17″). In particular, relation (2.17″) holds, since in the case when $w(R,T)$ satisfies case (b) of Theorem 2, then all non-trivial Fourier coefficients of $\bar\rho$ equal zero.

Let us also show that the Fourier coefficients in the second line of formula (2.20) corresponding to the function $zK(x)$ tend to zero as $z\to\infty$. This relation holds, because $B\neq 0$ in this case, and the Riemann lemma yields that

$$\begin{aligned}&\frac{1}{(b-a)}\int_a^b e^{izB(m_1,\dots,m_{2s})K(u)}\,du=\\ &\frac{1}{(b-a)}\int_{K_1(a)}^{K_1(b)} \frac{e^{izB(m_1,\dots,m_{2s})u}}{K'(K^{-1}(u))}\,du\to 0 \quad \text{as } z\to\infty.\end{aligned}\tag{2.21}$$

Let $g(u,v)$ be a bounded continuous function. We get similarly to the argument of Proposition 1 from relations (2.17) (2.17′) and (2.17″) that

$$
\begin{aligned}
&\lim_{T\to\infty} \frac{1}{(b-a)T} \int_{aT}^{bT} g(F(R), F(R+w(R,T))\, dR \\
(2.22)\qquad &= \lim_{T\to\infty} \frac{1}{(b-a)T} \int_{aT}^{bT} g(G(V(U(R))))\, dR \\
&= \lim_{T\to\infty} \int g(G(V(u))\bar\rho_{T,w,(a,b)}\,(du) = \int g(G(V(u))\bar\rho\,(du)
\end{aligned}
$$

with $\bar\rho = \bar\rho^{K(x)}_{(a,b)}$ if case (a) and $\bar\rho = \bar\rho^{\infty}_{(a,b)}$ if case (b) of Theorem 2 holds. Hence

$$
(2.23)\qquad \lim_{T\to\infty} \frac{1}{(b-a)T} \int_{aT}^{bT} g(F(R), F(R+w(R,T))\, dR = \int g(u,v)\bar\mu(du,dv)
$$

with

$$
(2.24)\qquad \bar\mu(\mathbf{A}) = \bar\rho\{u\colon\, u\in \mathcal{V}\times\mathcal{V},\ G(V(u))\in \mathbf{A}\}
$$

for all measurable sets $\mathbf{A}\in \mathbb{R}^2$.

Relation (2.23) implies the weak convergence of the measures $\mu_{T,w,(a,b)}$ to $\bar\mu$. Formula (2.24) together with the form of the measures $\bar\rho$ imply that the measure μ has the prescribed form if cases (a) of Theorem 2 holds, i.e. it depends only on the function $K(x)$. If case (b) of Theorem 2 holds, then a comparison of formulas (2.12) and (2.24) together with relation (2.17″) and the product form of the functions G and V in formulas (2.14) and (2.15) imply formula (1.8). Relation (1.9) can be deduced from the weak convergence in the same way as the analogous result in Proposition 1. To prove formula (1.10) and continuity of the measures $\mu^{zK(x)}_{(a,b)}$ (in the variable z) it is enough to prove the continuity of the expression $\int g(G(V(u)))\bar\rho^{zK(x)}_{(a,b)}(du)$. Because of the Weierstrass approximation theorem it is enough to check the continuity of the Fourier coefficients. This follows from relations (2.20) and (2.21). Proposition 2 is proved. □

Proof of Proposition 3. The proof is based on a representation similar to Proposition 2. Define the maps $U\colon \mathbb{R}^1 \to \mathcal{V}\times\mathcal{V}$

$$
(2.25)\qquad \begin{gathered} U(R) = U(R,x) = \\ \{R\tau_k \pmod 1,\ k=1,\dots,s\,,\ (R+x)\tau_k \pmod 1,\ k=1,\dots,s\} \end{gathered}
$$

and the maps $V(u_1,\dots,u_{2s})$ and $G(u_1,\dots,u_{2p})$ by formulas (2.14) and (2.15) as in the proof of Proposition 2. Introduce the measures

$$
(2.26)\qquad \hat\rho_{T,x}(\mathbf{A}) = \frac{1}{T}\lambda\{R\colon 0\le R\le T,\ U(R,x)\in \mathbf{A}\}
$$

for any measurable set $\mathbf{A} \subset \mathcal{V} \times \mathcal{V}$. Then

$$\hat{\rho}_{T,x} \Rightarrow \hat{\rho}^x \quad \text{as } T \to \infty \tag{2.27}$$

with Fourier coefficients
(2.28)

$$L^x(m_1, \dots, m_{2s}) = \begin{cases} 0 & \text{if } A(m_1, \dots, m_{2s}) \neq 0 \\ \exp\{iB(m_1, \dots, m_{2s})x\} & \text{if } A(m_1, \dots, m_{2s}) = 0 \end{cases}$$

with the functions $A(m_1, \dots, m_{2s})$ and $B(m_1, \dots, m_{2s})$ defined in (2.19′). Then we get the proof of the identity (1.14) as in the proof of Proposition 2 with the limit measure

$$\nu^x(\mathbf{A}) = \hat{\rho}^x\{u\colon u \in \mathcal{V} \times \mathcal{V},\ G(V(u)) \in \mathbf{A}\} \tag{2.29}$$

for all measurable $\mathbf{A} \in \mathbb{R}^2$. The expression at the right-hand side of (1.14) is clearly a bounded function, and it is a continuous function of x, because the Fourier coefficients of $\hat{\rho}^x$ are continuous functions of x. A comparison of the Fourier coefficients in (2.20) and (2.28) yields that

$$\hat{\rho}^{K(x)}_{(a,b)} = \frac{1}{(b-a)} \int_a^b \hat{\rho}^{K(x)}\, dx = \frac{z}{(b-a)} \int_{K(a)}^{K(b)} \frac{1}{K'K^{-1}(x)} \hat{\rho}^x\, dx .$$

This relation together with (2.24) and (2.29) imply relation (1.15). Proposition 3 is proved. □

3. Proof of the Theorem

First we formulate an estimate which enables us to generalize the results of Section 2 to functions satisfying (1.1) and (1.2).

Let $g(u,v)$ be a continuous function such that $|g(u,v)| < A(u^2+v^2)+B$ with some $A>0$ and $B>0$, $F(R)$ a function satisfying (1.1) and (1.2), $F_p(R)$ the trigonometrical series containing the first p terms of $F(R)$, and let some numbers $0<a<b\leqq 1$ and functions $w(R,T)$, satisfying either case (a) or case (b) of Theorem 2. We claim that for all $\varepsilon>0$ there are some thresholds $p_0 = p_0(\varepsilon)$ and $T_0 = T_0(p,\varepsilon)$ for $p > p_0$ such that

$$\left| \frac{1}{(b-a)T} \int_{aT}^{bT} g(F(R), F(R+w(R,T)))\, dR - \frac{1}{(b-a)T} \int_{aT}^{bT} g(F_p(R), F_p(R+w(R,T)))\, dR \right| < \varepsilon \tag{3.1}$$

for any $p > p_0$ and $T > T_0(p, \varepsilon)$. Moreover, the threshold p_0 can be chosen depending only on ε, a, b and $g(u,v)$, but not depending on the choice of the function $w(R,T)$ of which we only require that it satisfied the conditions of Theorem 2.

To prove relation (3.1) first we make the following observation: For any $(u,v) \in \mathbb{R}^2$ and $(u_0, v_0) \in \mathbb{R}^2$ and $1 > \eta > 0$ there exist some constants $K = K(\eta)$ and K_0 depending only on the function $g(u,v)$ such that

$$
\begin{aligned}
&|g(u,v) - g(u_0,v_0)| < \eta + K((u-u_0)^2 + (v-v_0)^2) \\
&+K_0(u_0^2+v_0^2)I(\{u_0^2+v_0^2 > \eta^{-1}\}) + K_0(u^2+v^2)I(\{u^2+v^2>\eta^{-1}\})\,,
\end{aligned}
\tag{3.2}
$$

where $I(A)$ denotes the indicator function of the set A.

Indeed, relation (3.2) holds with

$$
K_0 = 2 \sup_{u^2+v^2 \geqq 1} \frac{|A(u^2+v^2)+B|}{u^2+v^2} \leqq 2(A+B)
$$

if $u_0^2 + v_0^2 > \eta^{-1}$ or $u^2+v^2 > \eta^{-1}$. On the complementary set this inequality holds if $(u_0-u)^2 + (v_0 - v)^2 < \delta$ with some $\delta = \delta(\eta)$ because of the uniform continuity of the function $g(u,v)$ on this set. Finally, relation (3.2) holds on the remaining set if $K = K(\eta)$ is chosen sufficiently large. Relation (3.2) can be rewritten in a simpler form. We can apply the inequality

$$
(u^2+v^2)I(\{u^2+v^2>\eta^{-1}\}) \leqq 2u^2 I(\{u^2 > (2\eta)^{-1}\}) + 2v^2 I(\{v^2 > (2\eta)^{-1}\})\,,
$$

and write with the help of this relation that

$$
\begin{aligned}
|g(u,v) - g(u_0,v_0)| &< \eta + K((u-u_0)^2 + (v-v_0)^2) \\
&\quad + K_0[(u_0^2 I(\{u_0^2>\eta^{-1}\}) + v_0^2 I(\{v_0^2 > \eta^{-1}\})] \\
&\quad + K_0[(u^2 I(\{u^2>\eta^{-1}\}) + v^2 I(\{v^2 > \eta^{-1}\})]\,.
\end{aligned}
\tag{3.2$'$}
$$

with a new constant K which corresponds to the bound $\eta/2$ and with a new constant K_0 which is the double of the original one.

We shall prove (3.1) from (3.2′) with the choice $(u_0, v_0) = (F_p(R), F_p(R + w(R,T)))$, $(u,v) = (F(R), F(R+w(R,T)))$ with an appropriate $\eta > 0$ and $p = p(\eta)$ and then by integration with respect to R.

By relation (3.2′)

$$
\begin{aligned}
&|g(F(R), F(R+w(R,T))) - g(F_p(R), F_p(R+w(R,T)))| \\
&< \eta + K[(F(R)-F_p(R))^2 + (F(R+w(R,T)) - F_p(R+w(R,T)))^2] \\
&+K_0[F_p(R)^2 I\{F_p(R)^2>\eta^{-1}\} + F_p(R+w(R,T))^2 I\{F_p(R+w(R,T))^2>\eta^{-1}\}] \\
&+K_0[F(R)^2 I\{F(R)^2>\eta^{-1}\} + F(R+w(R,T))^2 I\{F(R+w(R,T))^2 > \eta^{-1}\}].
\end{aligned}
$$

The inequality

$$\left| \frac{K_0}{(b-a)T} \int\limits_{aT}^{bT} F(R)^2 I\{F(R)^2 > \eta^{-1}\}\, dR \right| < \frac{\varepsilon}{8} \tag{3.3}$$

holds, if $\eta < \eta(\varepsilon)$ and $T > T(\varepsilon)$. Indeed, by relation (1.5) there is some $\bar p = \bar p(\varepsilon)$, and $T(\varepsilon)$ in such a way that

$$\frac{K_0}{(b-a)T} \int\limits_{aT}^{bT} |F(R) - F_{\bar p}(R)|^2\, dR < \frac{\varepsilon}{32}$$

for $T > T(\varepsilon)$. The function $|F_{\bar p}(R)|$ is bounded. Put $\eta = \inf\limits_R 4|F_{\bar p}(R)|^{-2}$. Then the last inequality implies (3.3), since $F(R)^2 < 4|F(R) - F_{\bar p}(R)|^2$ on the set $\{F(R)^2 > \eta^{-1}\}$. We also claim that

$$\left| \frac{K_0}{(b-a)T} \int\limits_{aT}^{bT} F(R + w(R,T))^2 I\{F(R + w(R,T))^2 > \eta^{-1}\}\, dR \right| < \frac{\varepsilon}{8} \tag{3.3$'$}$$

if $\eta < \eta(\varepsilon)$ and $T > T(\varepsilon)$. This can be proved similarly to (3.3) with some modification. Apply the change of variables $u = R + w(R,T)$ in the integral in (3.3′). Since $w(R,T)$ satisfies the conditions Theorem 2, $\frac{du}{dR} \to 1$ uniformly for $aT \leqq R \leqq bT$, as $T \to \infty$. The domain of integration after this change of variable is the interval $[aT(1+o(1)), bT(1+o(1))]$. Hence after this change of variables the integral in (3.3′) can be estimated in the same way as in (3.3). Relations (3.3) and (3.3′) remain valid if the function $F(R)$ is replaced by $F_p(R)$ with $p > \bar p$, and $T > T(\varepsilon, p)$.

Choose η so that relations (3.3) and (3.3′) and their variants for the function $F_p(R)$ hold and $\eta < \varepsilon/4$. Then, because of relation (1.5) and the argument in the proof of (3.3′) some thresholds $p_0 = p(\eta)$ and $T_0 = T_0(\eta, p)$ can be chosen in such a way that for $p > p_0$ and $T > T_0$

$$\frac{K}{(b-a)T} \int\limits_{aT}^{bT} (F(R) - F_p(R))^2\, dR < \frac{\varepsilon}{8} \tag{3.4}$$

and

$$\frac{K}{(b-a)T} \int\limits_{aT}^{bT} (F(R + w(R,T)) - F_p(R + w(R,T)))^2\, dR < \frac{\varepsilon}{8} \tag{3.4$'$}$$

with the constant $K = K(\eta)$ appearing in formula (3.2′). Formulas (3.3), (3.3′), their variants for the function $F_p(R)$, (3.4) and (3.4′) together with the relation $\eta < \varepsilon/4$ imply (3.1).

It follows from (3.1) and relation (1.9) already proved for the function F_p that

$$\limsup_{T\to\infty} \left| \frac{1}{(b-a)T} \int_{aT}^{bT} g(F(R), F(R+w(R,T)))\, dR \right.$$
$$\left. - \int g(u,v)\bar{\mu}(F_p)(\,du,\, dv) \right| < \varepsilon$$

for all $p > p(\varepsilon)$ with $\bar{\mu}(F_p) = \mu_{(a,b)}^{K(x)}(F_p)$ if $w(R,T)$ satisfies the conditions of case (a) and with $\bar{\mu}(F_p) = \mu_{(a,b)}^{\infty}(F_p)$ if it satisfies the conditions of case (b) of Theorem 2. This relation implies that

$$\lim_{T\to\infty} \frac{1}{(b-a)T} \int_{aT}^{bT} g(F(R), F(R+w(R,T)))\, dR \tag{3.5}$$
$$= \lim_{T\to\infty} \int g(u,v)\mu_{T,w,(a,b)}(\,du,\, dv) = \lim_{p\to\infty} \int g(u,v)\, \mu(F_p)(du,\, dv)$$

with the same choice of the measure $\bar{\mu}(F_p)$ as in the previous formula. The last relation also means that all limits in this formula exist. Since relation (3.5) also holds for $g(u,v) = u^2 + v^2$, hence the measures $\mu_{T,w,(a,b)}$ are uniformly tight, and we get by applying relation (3.5) that there exists the limit

$$\bar{\mu} = \lim_{T\to\infty} \mu_{T,w,(a,b)} = \lim_{p\to\infty} \bar{\mu}(F_p)\,, \tag{3.6}$$

with the same measures $\bar{\mu}(F_p)$ as in (3.5) and in the previous formula, and also relation (1.9) holds with the function $F(R)$. Moreover, for a fixed bounded continuous function $g(u,v)$ the limit

$$\lim_{p\to\infty} \int g(u,v)\, \mu_{(a,b)}^{zK(x)}(F_p)(du,\, dv) = \int g(u,v)\, \mu_{(a,b)}^{zK(x)}(du,\, dv)$$

is uniform in z, and this fact together with the continuity properties of the measures $\mu_{(a,b)}^{zK(x)}(F_p)$ imply the continuity of the measures $\mu_{a.b}^{zK(x)}$ for $0 < z < \infty$ and relation (1.10). This completes the proof of Theorem 2 with the exception of formula (1.8).

A similar, but simpler argument shows that relation (3.1) holds if the pair of functions $g(F(R), F(R+w(R,T)))$ and $g(F_p(R), F_p(R+w(R,T)))$ are replaced by the pairs of functions $g(F(R), F(R+x))$ and $g(F_p(R), F_p(R+x))$ or $g(F(R))$ and $g(F_p(R))$.

The same argument as in the proof of Theorem 2 with the first replacement yields the existence of the limit

$$\lim_{p\to\infty} \nu^x(F_p) = \nu^x$$

together with relation (1.14) and the continuity of the integral in (1.14) as a function of x. If $|g(u,v)| < A(u^2+v^2)+B$, then the bound

$$\left|\int g(u,v)\nu^x(\,du,\,dv)\right| \leqq \limsup_{T\to\infty} \frac{1}{T}\int_0^T A[(F(R)^2+F(R+x)^2)+B]\,dR \leqq \text{const.}$$

holds with a constant independent of x, as we claimed. The identity (1.15) follows by a simple limiting procedure when $F(R)$ is approximated by the functions $F_p(R)$. This completes the proof of Theorem 3.

The second replacement in formula (3.1) supplies the proof of Theorem 1 in the same way. Finally, since the measures $\mu(F_p)$ tends to μ and $\mu^{\infty}_{(a,b)}(F_p) = \mu(F_p)\times\mu(F_p)$ tend to $\mu^{\infty}_{(a,b)}$ as $p\to\infty$, a limiting procedure implies the second relation in (1.10). The proof of the Theorems is completed.

Proof of the Corollary. Let us first consider the case when a simple function $h(u) = h_k(u)$ is chosen which is the linear combination of the indicator function of certain intervals. In this case formula (1.15) implies (1.16). A general density function $h(u)$ can be approximated by a sequence $h_k(u)$ in such a way that

$$\lim_{k\to\infty} \int_a^b |h_k(u) - h(u)|^p\,du = 0\,, \qquad p = 1,\,2.$$

Then a simple limiting procedure $h_k \to h$ gives the proof of the Corollary. □

4. Some comments and generalizations

The limit distributions in Theorems 1, 2 and 3 were given as the limit of a sequence of probability measures $\mu(F_p)$, $\mu^{K(x)}_{(a.b)}(F_p)$, $\mu^{\infty}_{(a,b)}$ and $\nu^x(F_p)$ which appeared as the solution of the corresponding problems when the function F was replaced by finite trigonometrical series. To describe these approximating measures we had to express the frequencies $\lambda_1,\dots,\lambda_p$ as the linear combination of some numbers $\tau_1,\dots,\tau_s$ linearly independent over the rational numbers with integer coefficients. This is possible for all finite subsets of the frequencies λ_n appearing in (1.1), but may be not possible for all λ_n simultaneously. We shall say that the function $F(R)$ has almost independent frequencies if all frequencies $\{\lambda_n,\ n=1,2,\dots\}$ in formula (1.1) can be expressed simultaneously as the finite linear combination of some numbers $\tau_1,\ \tau_2,\dots$ linearly independent over the rational numbers with integer coefficients. In this case the limit distributions in Theorems 1 and 3 can be described directly. If the function $F(R)$ arises as the Fourier expansion of a randomly magnified convex domain with a nice boundary, then it has almost periodic frequencies in the generic case, but not always. The case when it has almost independent frequencies is discussed in detail in paper [3]. This property holds for instance if the function $F(R)$ gives the Fourier expansion of the number of lattice points in concentrical circles of radius R.

If the function $F(R)$ in (1.1) is a finite trigonometrical series, then the limit distributions appearing in Theorems 1 and 3 have a relatively simple form. They are the distribution of a random variable of the form $\sum a_j e^{2\pi i T_j}$, where all (finitely many) T_j are linear combinations of independent on the interval [0, 1] uniformly distributed random variables with integer coefficients.

This can be seen by following the construction of the limit measures in the proofs of Section 2. Indeed, to understand the structure of the limit measure μ appearing in Theorem 1 let us express the frequencies λ_n in the form (2.3), and define the functions V and G by means of this formula as it was done in (2.6) and (2.7). Then formula (2.12) states that the measure μ is equal to the distribution of the random variable $G(V(\xi))$, where $\xi=(\xi_1,\dots,\xi_s)$ is a uniformly distributed random variable on the torus $\mathcal{V}$ defined in vector (2.4). The coordinates of the random vector $V(\xi_1,\dots,\xi_s)$ are linear combinations of independent, uniformly distributed random variables in [0, 1] with integer coefficients, and this fact together with the form of the function G gives a representation of $G(V(\xi))$ in the above described form.

The measures ν^x can also be represented in a similar way. Here again, the measure ν^x is the limit distribution of the random variable $G(V(\xi))$, but now the functions G and V are defined in (2.14) and (2.15), and $\xi=(\xi_1,\dots,\xi_{2s})$ is a $\bar\rho^x$ distributed random vector, where $\bar\rho^x$ is the probability measure on $\mathcal{V}\times\mathcal{V}$ with Fourier coefficients (2.28). Actually a $\bar\rho^x$ distributed random vector has a very simple representation. Indeed, let $\eta_1,\dots,\eta_s$ be independent uniformly distributed random variables on the unit interval [0, 1], and let $\eta_{s+k}=\eta_k+$ $+\tau_k$, (mod 1), $k=1,\dots,s$. Then relation (2.19′) and the expression for the Fourier coefficients (2.28) imply that $(\eta_1,\dots,\eta_{2s})$ is a $\bar\rho^x$ distributed random vector. Then, since the vectors $(\eta_k,\eta_{k+s})=(\eta_k,\eta_k+\tau_k \pmod 1)$ are independent, and η_k is uniformly distributed in [0, 1], the same argument works as in the case of the measure μ.

If the function $F(R)$ has almost independent frequencies, then the set of frequencies $\{\lambda_n,\ n=1,\dots,p\}$ can be expressed in (2.3) with numbers τ_k and coefficients $A(n,k)$ independent of p. In the polynomials whose distribution equal $\mu(F_p)$ and $\nu^x(F_p)$ the same independent random variables can be used for different p. Then the limit distribution μ and ν^x are the distribution of the limit of the random variables constructed for the representation of $\mu(F_p)$ and $\nu^x(F_p)$. Let us observe that the random variables constructed in such a way converge in L_2 norm as $p\to\infty$, and not only their distribution is convergent. This convergence holds, because of (1.2) and the orthogonality of the terms $e^{2\pi i T_j}$ appearing in these expressions. (Actually this representation could be proved by working directly with the function $F(R)$ instead of its approximation by the functions $F_p(R)$.)

In certain cases the above representation is even simpler. So e.g. if $F(R)$ is the Fourier expansion of the number of lattice points in a circle of radius R, then $F(R)$ has almost independent periods. Moreover, each λ_n can be expressed as a single τ_k multiplied by an integer. In this case the above argument yields a representation of μ and ν^x as the distribution of sums of independent random variables. The measures $\mu^{K(x)}_{(a,b)}$ appearing in Theorem 2 do not have such a simple representation as μ or ν^x. On the other hand, they can be expressed as the mixture of the measures ν^x as it is done in (1.15). This relation together with the continuity of the measures ν^x also implies that

$$\lim_{b\to a}\mu^{K(x)}_{(a,b)}=\nu^{K(a)}\,.$$

Theorems 2 and 3 can be generalized in a natural way. The vectors

$$\left(F\Bigl(R+\sum_{j=1}^{l}w_j(R,T)\Bigr),\quad l=1,\ldots,m\right),\qquad aT<R<bT$$

or

$$\left(F\Bigl(R+\sum_{j=1}^{l}x_j\Bigr),\quad l=1,\ldots,m\right),\qquad 0<R<T$$

have a limit distribution as $T\to\infty$ if all $w_j(R,T)$ satisfy the conditions of Theorem 2. They also have the continuity properties analogous to Theorem 2 and 3. In particular, the limit of the first vector equals the m-fold direct product of the measure μ if all $w_j(R,T)$ satisfy the conditions of case (a) in Theorem 2 with functions $z_jK(x)$, and $z_j\to\infty$. The proofs can be done by slightly modifying the method of the present paper. We omit the details.

In Theorem 2 we assumed that $aT\leqq R\leqq bT$ with some $a>0$. Some of the results follow automatically also for $a=0$ from our results, but to generalize all statements of Theorem 2 to the case when the parameter a can take also the value zero some additional conditions must be imposed. To carry out all required limiting procedures we must know that

$$\lim_{\varepsilon\to 0}\limsup_{T\to\infty}\frac{1}{T}\int_0^{\varepsilon T}F(R+w(R,T))\,dR=0$$

To guarantee the last relation some additional properties should be imposed on the function $w(R,T)$. Since the restriction $a>0$ is not essential in applications we proved Theorem 2 only under the condition $a>0$.

The results of this paper were proved originally for finite trigonometrical sums in Section 2, and then in Section 3 these results were generalized to functions which can be well approximated by finite trigonometrical sums. The content of formulas (1.1) and (1.2) was the possibility of such a good approximation. In applications this condition can be checked. On the other hand, the weak convergence of the random variables $F(R)$, $F((R),F(R+ +w(R,T)))$ or $(F(R),F(R+x))$ in Theorems 1, 2 and 3 also hold if formulas (1.1) and (1.2) are replaced by the following weaker condition: There exists a sequence of finite trigonometrical sums $F_p(R)$, $p=1,2,\dots$, such that

$$\lim_{p\to\infty}\limsup_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\min\{1,|F_p(R)-F(R)|\}\,dR=0\,. \tag{4.1}$$

Similar conditions were formulated in paper [2] or [4].

We only briefly explain why formula (4.1) implies the weak convergence in Theorems 1, 2 and 3. If we consider functions of the form $g(u,v)=g_{s,t}(u,v)= =e^{i(su+tv)}$, then one can show by means of condition (4.1) and the relation $\frac{\partial}{\partial R}(R+w(R,T))=1+o(1)$ that under the conditions of Theorem 2 the limits

$$\begin{aligned}&\lim_{T\to\infty}\frac{1}{(b-a)T}\int_{aT}^{bT}g(F(R),F(R+w(R,T)))\,dR\\ &=\lim_{T\to\infty}\lim_{p\to\infty}\frac{1}{(b-a)T}\int_{aT}^{bT}g(F_p(R),F_p(R+w(R,T)))\,dR\end{aligned} \tag{4.2}$$

exist. Hence to prove the weak convergence of the distribution of the random vectors $(g(F(R)),g(F(R+w(R,T))))$, $aR\leqq T\leqq bT$, it is enough to check the compactness of these distributions in the weak topology. To do this it is enough to show that for any $\varepsilon>0$ there exists a constant $K=K(\varepsilon)$ such that the following relation holds. The function $h(u,v)=h_K(u,v)=H_K(u^2+v^2)$, where $H_K(u)$ is defined by the relations $H_K(u)=0$ for $|u|\leqq K$, $H_K(u)=1$ for $|u|\geqq 2K$, and $H_K(u)$ is given by linear interpolation for $K<|u|<2K$, satisfies the inequality

$$\limsup_{T\to\infty}\frac{1}{(b-a)T}\int_{aT}^{bT}h(F_p(R),F_p(R+w(R,T)))\,du\,dv<\varepsilon\,.$$

To prove this relation, choose a number $\bar{p}$ such that for $T > T(p)$

$$\frac{1}{2T}\int_{-T}^{T} \min\{1, |F_{\bar{p}}(R) - F(R)|\}\, dR < \frac{(b-a)}{2}\varepsilon .$$

Then, since $F_{\bar{p}}(R)$ is a bounded function we can choose $K = 1 + 2\sup\limits_R F_{\bar{p}}(R)$. It is not difficult to see that $\int h(F_{\bar{p}}(R), F_{\bar{p}}(R + w(R,T)))\, dR = 0$, and relation (4.2) holds with this choice of the function $h_K(u,v)$. The analogue of Theorem 2 under condition (4.1) can be proved by working out the details. The modified version of Theorems 1 and 3 can be proved similarly.

In this paper we did not discuss such functions $w(R,T)$ which satisfy the relation

$$\lim_{T\to 0} \sup_{0\le R\le T} w(R,T) = 0 .$$

The reason for this omission is not our disinterest for this case. Actually, the description of this case is a very exciting problem. This is related to the investigation of the limit behaviour of the number of lattice points in randomly chosen thin strips. This is a very interesting problem with many unsolved conjectures and few rigorous results. The methods of the present paper are not sufficient to study such problems. Here some essentially new ideas are needed.

REFERENCES

[1] Besicovitch, A. S., *Almost periodic functions*, Dover Publications, New York, 1955. *MR* **16**-817

[2] Bleher, P. M., On the distribution of the number of lattice points inside a family of convex ovals, *Duke Math. J.* **67** (1992), 461–481. *MR* **93h**:11110

[3] Bleher, P. M. and Lebowitz, J. L., Energy-level statistics of model quantum systems: Universality and scaling in a lattice-point problem, *J. Statist. Phys.* (to appear).

[4] Heath-Brown, D. R., The distribution and moments of the error term in the Dirichlet divisor problem, *Acta Arith.* **60** (1992), 389–415. *MR* **93e**:11114

(Received August 25, 1994)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H-1364 BUDAPEST
HUNGARY

and

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
BOLYAI KOLLÉGIUMA
AMERIKAI ÚT 96
H-1145 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 225–236

LARGE DEVIATION RESULTS FOR COVER TIMES

T. F. MÓRI

Dedicated to Pál Révész for his sixtieth birthday

Abstract

Considering an infinite string of i.i.d. random letters drawn from a finite alphabet we define the cover time W_n as the number of random letters needed until each pattern of length n appears at least once as a substring. In this paper we derive large deviation type results for the asymptotic distribution of W_n.

1. Introduction and expository results

Let $\mathbf{X}$ be finite alphabet of size d ($d \geqq 2$), we can suppose $\mathbf{X} = \{1, 2, \dots, d\}$. Elements of $\mathbf{X}^n$, $n \geqq 1$ will be called *patterns* or *words* of length n. Let $\{p_1, p_2, \dots, p_d\}$ be a probability distribution on $\mathbf{X}$. Without loss of generality we can assume that all probabilities are positive and they are in increasing order. Let p and q denote the smallest and the second smallest (different) values among the probabilities, resp., and let a and b be the corresponding multiplicities. Thus $p_1 = \dots = p_a = p$, $p_{a+1} = \dots = p_{a+b} = q > p$, if $a < d$, and $p_{a+b+1} > q$, if $a + b < d$. Consider an infinite sequence of i. i. d. random letters $X_1, X_2, \dots$ drawn from $\mathbf{X}$ according to the given distribution. For any pattern $A = (a_1 \dots a_n) \in \mathbf{X}^n$ define $\mathbf{P}(A) = \prod_{i=1}^n p_{a_i}$, and

$$T(A) = \inf\{N \geqq n : (X_{N-n+1} X_{N-n+2} \dots X_N) \equiv A\}$$

the *waiting time* until A is observed as a substring in the sequence of experiments. Finally, for any subset $H_n \subset \mathbf{X}^n$ define

$$W(H_n) = \max\{T(A) : A \in H_n\};$$

the number of random letters needed until each pattern of H_n appears at least once. This random variable is called the *cover time* associated with H_n. In the most interesting particular case, when $H_n = \mathbf{X}^n$, the notation $W_n = W(\mathbf{X}^n)$ will be used.

Several papers have been devoted to the study of weak and a. s. asymptotic properties of cover times, most of them in the symmetric case, i. e., where the random letters are drawn according to the uniform distribution

1991 *Mathematics Subject Classification*. Primary 60E10.

Key words and phrases. Pattern, run, waiting time, cover time, large deviations.

0081–6906/96/\$ 5.00 ©1996 *Akadémiai Kiadó, Budapest*

on $\mathbf{X}$, see [1], [2], [5], [6]. The limit distribution in the symmetric case was found to be of Gumbel type [5], in addition, the speed of convergence (measured by the uniform distance between distribution functions) appeared surprisingly fast. In the most interesting case, where H_n is comparable with $\mathbf{X}^n$ in size, the rate of convergence proved to be exponential, which made it possible to derive strong results on the a. s. behaviour of cover times [6].

PROPOSITION 1. [6]. *Suppose every letter has the same probability to be drawn (symmetric case). Let* $F(t)=\exp(-e^{-t})$, *then*

$$\sup_t \left|\mathbf{P}(d^{-n}W(H_n)-\log|H_n|\leqq t)-F(t)\right| = o\left(|H_n|^{-\frac{1}{6}\frac{d-1}{d+1}}\right)$$

as n and $|H_n|$ tend to infinity.

For sake of simplicity, in the non-symmetric case $W_n=W(\mathbf{X}^n)$ was only considered. The case $a>1$ has turned out to be very similar to the symmetric case, because, roughly speaking, words that are built up entirely from the least probable letters take the longest time to appear in the covering process. On the contrary, the case $a=1$ is more interesting, for the limit distribution is not always of pure Gumbel type, and besides, the additive normalizing constant is proportional to $\log n$ instead of n. In order to formulate the limit theorem, let us first define $\lambda=q/p>1$, and $k=\left\lceil\frac{\lambda}{\lambda-1}\right\rceil=\left\lceil\frac{q}{q-p}\right\rceil$. This k maximizes the sequence $\{i\lambda^{-i}, i\geqq 0\}$.

PROPOSITION 2. [7]. (a) *Suppose $a>1$ and let $Y_n=p^nW_n-\log a^n$. Then*

$$\lim_{n\to\infty}\mathbf{P}(Y_n\leqq t)=F(t).$$

(b) *Suppose $a=1$ and let $Y_n=p^n\lambda^kW_n-\log\left(b^k\binom{n}{k}\right)$. Then*

$$\lim_{n\to\infty}\mathbf{P}(Y_n\leqq t)=F(t)\quad \textit{if}\quad \frac{\lambda}{\lambda-1}\neq k,$$

and

$$\begin{aligned}\lim_{n\to\infty}\mathbf{P}(Y_n\leqq t)&=\exp\left\{-e^{-t}-k(k!)^{-1/k}e^{-\frac{k-1}{k}t}\right\}\\&=F(t)F\left(\frac{k-1}{k}t+\log\left(k^{-1}(k!)^{1/k}\right)\right),\quad \textit{if}\quad \frac{\lambda}{\lambda-1}=k.\end{aligned}$$

As to the rate of convergence, it can be shown exponential in part (a), similarly to the symmetric case, but in part (b) the convergence is much slower: in fact, it is polynomial. Hence it would be of interest to derive non-uniform bounds. Such a result would be particularly important in the case $a=1$, where no useful estimates are known. The aim of the present note is to find large deviation type results in the asymptotic distribution theory of cover times.

In the proofs we shall often make use of the following basic facts concerning waiting times. The so-called *correlation of patterns* was introduced to measure the overlap between words. For arbitrary patterns $A=(a_1a_2\dots a_n)$ and $B=(b_1b_2\dots b_m)$ it is defined as

$$A*B=\sum_{i=1}^{n}\varepsilon_i\mathbf{P}(b_1\dots b_i)^{-1},$$

where $\varepsilon_i=1$ if $i\leq m$ and $(a_{n-i+1}\dots a_n)\equiv(b_1\dots b_i)$, otherwise $\varepsilon_i=0$. Expectation of minima of waiting times can be expressed in terms of correlations, e. g. $\mathbf{E}(T(A))=A*A$, and

$$\mathbf{E}(\min\{T(A),T(B)\})=\frac{A*A\cdot B*B-A*B\cdot B*A}{A*A+B*B-A*B-B*A} \tag{1.1}$$

(see [3]). These minima are approximately of exponential distribution. More precisely, the following estimation holds.

PROPOSITION 3. [4]. *Let* $A_1,A_2,\dots,A_r\in\mathbf{X}^n$ *and* $V=\min\{T(A_i):1\leqq i\leqq r\}$. *Suppose* $c=:2n\sum_{i=1}^{r}\mathbf{P}(A_i)<0.2$. *Then for every positive* y

$$\exp\{-(1+c)y\}\leqq\mathbf{P}(V>\mathbf{E}(V)y)\leqq\exp(c-y).$$

2. Large deviations

In this section we are going to formulate our main results. As we shall see, for the upper tail it is not very difficult to obtain large deviation-type results, but in the case of the lower tail all effort failed, although such a result would serve as a key for many actual unsolved problems in the theory of cover times.

THEOREM 1 *(Symmetric case). Suppose* $a>1$, *i. e., the smallest probability is multiple. Let* $H_n\subset\mathbf{X}^n$ *contain only patterns built up exclusively from letters of the smallest probability. Suppose* $|H_n|\to\infty$ *and* $t_n\to\infty$. *Let*

$$\alpha_n=\min\{1/(p^nA*A):A\in H_n\},$$

and assume that $\lim_{n\to\infty}\alpha_n=\alpha_\infty$.

(a) *If* $\alpha_\infty=1$, *and*

$$t_n\leqq\min\left\{\log|H_n|\left(\frac{2}{3(1-\alpha_n)}-1\right);|H_n|^{1/3}-\log|H_n|\right\},$$

then

$$\mathbf{P}(Y(H_n) > t_n) \asymp \exp(-t_n)$$

(that is, the ratio of the two sides is kept bounded away from zero and infinity).

(b) *If* $\alpha_\infty < 1$ *and* $np^n t_n \to 0$, *then*

$$\mathbf{P}(Y(H_n) > t_n) \asymp \exp\{-\alpha_\infty(t_n + \log|H_n|)\} + \exp(-t_n).$$

Note that this follows from Proposition 1 only when $\limsup\limits_{n\to\infty} t_n/\log|H_n| < \frac{1}{6}\frac{d-1}{d+1}$.

By Proposition 1 we have

$$\lim_{n\to\infty} \frac{p^n W(H_n)}{\log|H_n|} = 1$$

with probability 1. This limit theorem can easily be extended to a large deviation-type result by applying our Theorem 1.

COROLLARY. *If* $\min\{1/(p^n A * A) : A \in H_n\} \to \alpha_\infty$, *then for arbitrary fixed* $t \geqq 1$ *we have*

$$\lim_{n\to\infty} \frac{1}{\log|H_n|} \log \mathbf{P}\left(\frac{p^n W(H_n)}{\log|H_n|} > t\right) = -(t-1) \wedge \alpha_\infty t.$$

When $\alpha_\infty < 1$, the expression on the right-hand side is a piecewise linear function of t: the breakpoint is $t^* = 1/(1-a_\infty)$. For $t > t^*$ the distribution of $W(H_n)$ is determined mainly by waiting times for extremal words; they are significantly longer than those of typical words. For smaller values of t the effect of extremal words is not felt because of the large number of typical words.

For example, consider the classical case, when the random letters are uniformly distributed in $\mathbf{X}$, (i. e., $a = d$, $p = 1/d$), and the waiting time W_n until each word of length n appears at least once. Then

$$\lim_{n\to\infty} \frac{1}{n} \log_d \mathbf{P}\left(\frac{W_n}{d^n \log d^n} > t\right) = \begin{cases} -(t-1) & \text{if } 1 \leqq t < d, \\ -\dfrac{d-1}{d}t, & \text{if } d \geqq t. \end{cases}$$

In the non-symmetric case we only deal with the whole set $H_n = \mathbf{X}^n$, because, contrary to the symmetric case, the rate function of large deviation probabilities heavily depends on the structure of H_n, and lengthy discussions would conceal the point. So let $W_n = W(\mathbf{X}^n)$ again. From Proposition 2 it is clear that $\lim\limits_{n\to\infty} \frac{p^n W_n}{\log a^n} = 1$ in probability, when $a > 1$, and $\lim\limits_{n\to\infty} \frac{p^n W_n}{\log n} = \frac{k}{\lambda^k}$, when $a = 1$.

THEOREM 2 *(Non-symmetric case).* (a) *Suppose* $a > 1$, *i. e., the smallest probability is multiple. Then*

$$\lim_{n\to\infty} \frac{1}{n} \log_a \mathbf{P}\left(\frac{p^n W_n}{\log a^n} > t\right) = \begin{cases} -(t-1) & \text{if } 1 \leqq t < 1/p, \\ -(1-p)t & \text{if } 1/p \leqq t. \end{cases}$$

(b) *Suppose* $a = 1$, *i. e., the smallest probability is single. Define*

$$f_0(t) = (1-p)t, \quad f_i(t) = \lambda^i t - i, \quad i = 1, 2, \dots,$$

then

$$\lim_{n\to\infty} \frac{1}{\log n} \log \mathbf{P}\left(\frac{p^n W_n}{\log n} > t\right) = -\min\{f_i(t) : 0 \leqq i \leqq k\}, \quad t \geqq k\lambda^{-k}.$$

REMARK. Since the minimum of linear function is always concave, and the slope of f_i is increasing for $i = 1, 2, \dots$, in part (b) it is easy to see that

$$\min\{f_i(t) : 0 \leqq i \leqq k\} = f_i(t) \text{ if } t_{i+1} \leqq t < t_i,$$

where $t_0 = +\infty$, $t_i = \frac{\lambda}{\lambda-1}\lambda^{-i}$ for $1 \leqq i \leqq k$, and $t_{k+1} = k\lambda^{-k}$. The graph of this function is a polygon with vertices $(t_{k+1}, 0)$ and $\left(t_i, \frac{\lambda}{\lambda-1} - i\right)$, $1 \leqq i \leqq k$. Since the slope of f_0 is smaller than that of f_1, and $f_0(t_{k+1}) > 0$, $f_0(t_1) = \frac{1-p}{\lambda-1} < \frac{\lambda}{\lambda-1} - 1$, the straight line f_0 intersects this polygon at a unique point $(t^*; (1-p)t^*)$; thus $\min\{f_i(t) : 0 \leqq i \leqq k\} = f_0(t)$ if $t \geqq t^*$. Elementary calculation shows that $t^* = \frac{j}{\lambda^j - 1 + p}$ (this lies between t_{j+1} and t_j), where

$$j = \max\{i : f_i(t_i) > f_0(t_i)\} = \max\left\{i : \frac{i}{\lambda^i - 1 + p} < \frac{1}{1-p}\left(\frac{\lambda}{\lambda-1} - i\right)\right\}.$$

3. Proofs

PROOF OF THEOREM 1. For the sake of simplicity let us suppress the notation of dependence on n in H_n, t_n, and α_n. Let $x = t + \log|H|$, $c = 2np^n$. Applying Proposition 3 we obtain

$$\mathbf{P}(Y(H) > t) = \mathbf{P}(W(H) > p^{-n}x) \leqq \sum_{A\in H} \mathbf{P}(T(A) > p^{-n}x)$$
$$\leqq e^c \sum_{A\in H} \exp\left\{-\frac{x}{p^n A * A}\right\}.$$

Suppose first that $(1-\alpha)x \leqq 1$. Then

$$\sum_{A\in H} \exp\left\{-\frac{x}{p^n A * A}\right\} \leqq |H| e^{-\alpha x} \leqq e^{-t+1}.$$

Now let $1 < (1-\alpha)x$. Then the sum on the right-hand side is to be divided into three parts. Let us call a word $A \in H$ *good* if $p^n A * A \leqq \frac{x+1}{x}$, *extremal* if $p^n A * A \geqq \frac{\alpha+1}{2\alpha}$, and *bad* otherwise. According to Lemma 3 of [6], the number of bad words is $\leqq 4x$, and there are at most $\frac{8\alpha}{1-\alpha}$ of extremal words. Therefore

$$\sum_{A\in H}^{\cdot} \exp\left\{-\frac{x}{p^n A * A}\right\} = \sum_{A \text{ good}} + \sum_{A \text{ bad}} + \sum_{A \text{ extremal}}$$
$$\leqq |H| \exp\left(-\frac{x^2}{x+1}\right) + 4x \exp\left(-\frac{2\alpha x}{\alpha+1}\right) + \frac{8\alpha}{\alpha+1}\exp(-\alpha x).$$

Here the first term is less than $|H|e^{-x+1} = e^{-t+1}$. The second term can be estimated by the third one, since

$$4x \exp\left(-\frac{2\alpha x}{\alpha+1}\right) = \frac{4}{1-\alpha}\exp(-\alpha x)\left((1-\alpha)x \exp\left\{-\frac{\alpha}{\alpha+1}(1-\alpha)x\right\}\right)$$
$$\leqq \frac{4\alpha}{1-\alpha}\exp(-\alpha x)\frac{\alpha+1}{\alpha^2 e} \leqq \frac{9\alpha}{1-\alpha}\exp(-\alpha x).$$

The third term can be treated in the following way. From the definition of $A * A$ it is easy to see that the possible values of α_∞ are 1, and $1 - p^k$, $k > 1$. When $\alpha_\infty < 1$, this means that all the extremal words are periodic, and with the same period, provided n is large enough. Thus $\alpha_n = \alpha_\infty + O(p^n)$. Hence

$$\frac{1}{1-\alpha}\exp(-\alpha x) = O(\exp(-\alpha_\infty x)).$$

When $\alpha_\infty = 1$, by supposition we have that $1 < (1-\alpha)x \leqq \frac{2}{3}\log|H|$ and $\log x \leqq \frac{1}{3}\log|H|$, therefore

$$\frac{1}{1-\alpha}\exp(-\alpha x) = e^{-t}\exp\{(1-\alpha)x - \log(1-\alpha)x + \log x - \log|H|\} < e^{-t}.$$

This proves the assertion of part (a) in one direction.

For the lower estimation of $\mathbf{P}(Y(H) > t)$ let us first suppose that $\alpha_\infty < 1$ and $\alpha_\infty x < t$. Let $A \in H$ such that $A * A$ is maximal, then $W(H) \geqq T(A)$, hence by Proposition 3 we get

$$\mathbf{P}(Y(H) > t) = \mathbf{P}(W(H) > p^{-n}x) \geqq \mathbf{P}(T(A) > p^{-n}x) \geqq \exp\left\{-\frac{1+c}{p^n A * A}x\right\}$$
$$= \exp(-\alpha x)(1+o(1)) = \exp(-\alpha_\infty x)(1+o(1)) \geqq e^{-t}(1+o(1)).$$

Secondly, let α_∞ be arbitrary and $\alpha_\infty x \geqq t$. Define $\delta = |H|^{-1/3}$, and call a word $A \in H$ *good* if $A * A \leqq (1+\delta^2)p^{-n}$. Let H' be the subset of good words in H, then by Lemma 3 of [6] we have $|H| \geqq |H'| \geqq |H| - 4\delta^{-2}$, hence $|H'| \sim |H|$. Let us call a pair (A, B) from H' *good* if $A * B, B * A \leqq 2\delta p^{-n}$, *extremal* if either $A * B$ or $B * A$ is maximal, that is, equal to $p^{-n+1} + \cdots + p^{-1}$, and *bad* otherwise. As it was shown in [7], at most one of the correlations $A * B, B * A$ can be greater than $2\delta p^{-n}$. By analyzing (1.1) it is easy to see that

$$\frac{1}{p^n \mathbf{E}(T(A) \wedge T(B))} \geqq \begin{cases} 2(1 - 4\delta) & \text{for good pairs,} \\ \left(2 - \dfrac{p}{1-p}\right)(1 - 4\delta) & \text{for extremal pairs,} \\ \left(2 - \dfrac{p}{1-p^2}\right)(1 - 4\delta) & \text{for bad pairs,} \end{cases}$$

where $\wedge$ stands for minimum. We are going to show that

$$\mathbf{P}(W(H') > p^{-n}x) \geqq e^{-t}(1 + o(1)) \geqq \exp(-\alpha_\infty x)(1 + o(1)).$$

Let us start from the Bonferroni inequality

$$\mathbf{P}(W(H') > d^n x) \geqq S - Q$$

where

$$S = \sum_{A \in H'} \mathbf{P}(T(A) > p^{-n}x), \qquad Q = \sum_{\substack{A, B \in H' \\ A \neq B}} \mathbf{P}(T(A) \wedge T(B) > p^{-n}x).$$

By Proposition 3 again

$$S \geqq |H'| \exp\{-(1 + c)x\} = e^{-t}(1 + o(1)).$$

The sum of Q is divided into three parts according as the pair (A, B) is good, extremal or bad.

For good pairs we can apply Proposition 3, together with the inequality $\delta x \leqq \delta |H|^{1/3} \leqq 1$ (this is satisfied in the case $\alpha_\infty < 1$, too). We obtain that

$$\mathbf{P}(T(A) \wedge T(B) > p^{-n}x) \leqq \exp\{2c - 2x(1 - 4\delta)\} = O(e^{-2x}),$$

hence

$$\sum_{(A,B) \text{ good}} = O\left(|H'|^2 e^{-2x}\right) = O(e^{-2t}) = o(e^{-t}).$$

There are at most a^3 of extremal pairs. By Proposition 3 it follows that

$$\begin{aligned} \mathbf{P}(T(A) \wedge T(B) > p^{-n}x) &\leqq \exp\left\{2c - x\left(2 - \frac{p}{1-p}\right)(1 - 4\delta)\right\} \\ &= O\left(\exp\left\{-\frac{2 - 3p}{1 - p}x\right\}\right) = O(e^{-x}), \end{aligned}$$

thus the contribution of extremal pairs is $o(e^{-t})$.

Finally, bad pairs can be treated in the same way: they satisfy

$$\mathbf{P}(T(A)\wedge T(B)>d^n x)\leqq \exp\left\{2c-x\left(2-\frac{p}{1-p^2}\right)(1-4\delta)\right\}=O(e^{-4x/3}),$$

and since their number does not exceed $4|H|/\delta=4|H|^{4/3}$ (again by Lemma 3 of [6]), we have

$$\sum_{(A,B)\ \mathrm{bad}}=O\left(|H|^{4/3}e^{-4x/3}\right)=O(e^{-4t/3})=o(e^{-t}).$$

From all these we can see that

$$\mathbf{P}(Y(H)>t)\geqq e^{-t}(1+o(1))\geqq \exp(-\alpha_\infty x)(1+o(1)).$$

Thus the proof of the opposite direction is completed. □

Proof of Theorem 2 (b). Since the waiting time for a pattern A is more or less proportional to $\mathbf{P}(A)$, it is plausible to focus on patterns with only few, say i, letters of probability greater than p. Thus, for $i\geqq 0$ let

$$H(i)=\{A\in \mathbf{X}^n: \text{the number of letters with probability greater than } p \text{ is equal to } i\}$$

and

$$S_i=\sum_{A\in H(i)}\mathbf{P}\left(\frac{p^nT(A)}{\log n}>t\right)=\sum_{A\in H(i)}\mathbf{P}\left(\frac{T(A)}{A*A}>\frac{t\log n}{p^nA*A}\right),$$

then clearly

$$\mathbf{P}\left(\frac{p^nW_n}{\log n}>t\right)\leqq\sum_{i=0}^{n}\mathbf{P}\left(\frac{p^nW(H(i))}{\log n}>t\right)\leqq\sum_{i=0}^{n}S_i. \tag{3.1}$$

Firstly, let $i=0$. $H(0)$ consists of a single pattern for which $A*A\sim \sim\frac{1}{p^n(1-p)}$, hence by Proposition 3 we have

$$S_n\leqq 2\exp\{-t(1-p)(\log n)(1+o(1))\}=\exp\{-(f_0(t)+o(1))\log n\}. \tag{3.2}$$

Now let $1\leqq i<mk$, $i\neq k$, where $m\geqq 2/p>4$, integer. Then it is easy to see (Lemma 3.1 of [7]) that $\mathbf{P}(A)A*A=1+o(1)$ uniformly in i and A. Hence Proposition 3 implies

$$\begin{aligned}S_i&=\sum_{A\in H(i)}\mathbf{P}\left(\frac{T(A)}{A*A}>t\lambda^i(\log n)(1+o(1))\right)\\&\leqq(d-1)^i\binom{n}{i}\cdot 2\exp\{-t\lambda^i(\log n)(1+o(1))\}\\&=\exp\{-(\lambda^i t-i)(\log n)(1+o(1))\}=\exp\{-(f_i(t)+o(1))\log n\}.\end{aligned}\tag{3.3}$$

Finally, let $i \geqq mk$. Since $A * A \leqq \frac{1}{p\mathbf{P}(A)} \leqq p^{n-1}\lambda^{-i}$ holds obviously and $k > \frac{1}{\lambda - 1}$, for sufficiently large n we have

$$S_i \leqq (d-1)^i \binom{n}{i} \cdot 2\exp\{-pt\lambda^i \log n\} \leqq 2\frac{d^i}{i!}\exp\{-(p\lambda^i t - i)\log n\}$$
$$\leqq 2\frac{d^i}{i!}\exp\{-(p\lambda^i - \frac{i}{k}\lambda^k)\log n\}.$$

We are going to show that $p\lambda^i - \frac{i}{k}\lambda^k > 1$. Using that $\lambda^k = 1 + k(\lambda - 1) > 2$, $i\lambda^{-i} \geqq mk\lambda^{-mk}$, and $2^{m-1} \geqq (m-1)^2$, we get

$$p\lambda^i - \frac{i}{k}\lambda^k > \frac{i}{mk}p\lambda^{mk} - \frac{i}{k}\lambda^k = \frac{i}{mk}\lambda^k(p\lambda^{(m-1)k} - m) > 2(p2^{m-1} - m)$$
$$> 2\left(p\left(\frac{2}{p} - 1\right)^2 - \frac{2}{p}\right) = 2\left(\frac{2}{p} - 4 + p\right) \geqq 1.$$

Thus

$$\sum_{i=mk}^{n} S_i \leqq 2e^d \exp(-t\log n) \leqq \exp\{-(f_0(t) + o(1))\log n\}. \tag{3.4}$$

From (3.1)–(3.4) it immediately follows that

$$\limsup_{n\to\infty} \frac{1}{\log n} \log \mathbf{P}\left(\frac{p^n W_n}{\log n} > t\right) \leqq -\min\{f_i(t) : 0 \leqq i < mk\}.$$

On the right-hand side all f_i with $i > k$ can be left out of consideration, because

$$f_i(t) = \lambda^i(t - i\lambda^{-i}) > \lambda^k(t - t_{k+1}) = f_k(t) \text{ for } t \geqq t_{k+1}.$$

For the lower estimate let us define

$$H'(i) = \{A \in H(i) : \text{all letters of } A \text{ are of probability } p \text{ or } q\},$$

then clearly

$$\mathbf{P}\left(\frac{p^n W_n}{\log n} > t\right) \geqq \mathbf{P}\left(\frac{p^n W(H'(i))}{\log n} > t\right) \geqq S_i' - Q_i' \tag{3.5}$$

for every $i = 0, 1, \ldots, k$, where

$$S_i' = \sum_{A \in H'(i)} \mathbf{P}\left(\frac{p^n T(A)}{\log n} > t\right) = \sum_{A \in H'(i)} \mathbf{P}\left(\frac{T(A)}{A * A} > \frac{t\log n}{p^n A * A}\right),$$

and

$$Q'_i = \sum_{\substack{A,B\in H'(i)\\ A\neq B}} \mathbf{P}(T(A)\wedge T(B) > p^{-n}(\log n)t).$$

By Proposition 3, similarly to (3.2) we can write

(3.6) $S'_0 = S_0 \geqq \exp\{-t(1-p)(\log n)(1+o(1))\} = \exp\{-f_0(t)+o(1))\log n\}$,

and, obviously, Q'_0 is void.

When $1 \leqq i \leqq k$, similarly to the derivation of (3.3), Proposition 3 implies

$$S'_i = \sum_{A\in H'(i)} \mathbf{P}\left(\frac{T(A)}{A*A} > t\lambda^i(\log n)(1+o(1))\right)$$

$$\geqq b^i\binom{n}{i}\exp\{-t\lambda^i(\log n)(1+o(1))\} = \exp\{-(f_i(t)+o(1))\log n\}.$$

Let us turn to the estimation of Q'_i. Following the ideas of [7], we shall call a pair $(A, B\in H'(i), A\neq B)$ *distant* if the length of the maximal overlap between A and B in both directions is less than $n-w$, where $w\to\infty$, $\log w = o(\log n)$ is a fixed sequence of integers; *extremal* if $i=1$ and A overlaps B or B overlaps A in length $n-1$ (then one of A and B ends, the other one starts with a letter of probability q), and *close* otherwise. The following simple assertions are borrowed from [7].

The number of close pairs is not larger than

$$2b^{2i}w\binom{w}{i}\binom{n}{i} = \exp\{i(\log n)(1+o(1))\};$$

The number of extremal pairs is at most $2b^2$.

$$\frac{1}{p^n\mathbf{E}(T(A)\wedge T(B))} \geqq \begin{cases} \lambda^i(1+p)(1+o(1)) & \text{for close pairs,}\\ \lambda(1+o(1)) & \text{for extremal pairs,}\\ 2\lambda^i(1+o(1)) & \text{for distant pairs.}\end{cases}$$

Combining these with Proposition 3 we obtain

$$\sum_{(A,B)\text{ close}} \mathbf{P}(T(A)\wedge T(B) > p^{-n}(\log n)t)$$

$$\leqq \exp\{i(\log n)(1+o(1)) - \lambda^i(1+p)t(\log n)(1+o(1))\}$$

$$= \exp\{-(f_i(t)+\lambda^i pt + o(1))\log n\},$$

$$\sum_{(A,B)\text{ distant}} \mathbf{P}(T(A)\wedge T(B) > p^{-n}(\log n)t)$$

$$\leqq 2\binom{n}{i}^2 b^{2i}\exp\{-2\lambda^i t(\log n)(1+o(1))\}$$

$$= \exp\{-(2f_i(t)+o(1))\log n\},$$

and in the case $i = 1$

$$\sum_{(A,B)\text{ extremal}} \mathbf{P}(T(A) \wedge T(B) > p^{-n}(\log n)t) \leqq 4b^2 \exp\{\lambda t(\log n)(1+o(1))\}$$
$$= \exp\{-(f_i(t) + 1 + o(1))\log n\}.$$

Thus

$$S_i' - Q_i' \geqq \exp\{-(f_i(t) + o(1))\log n\}. \tag{3.7}$$

Now, (3.5)–(3.7) imply that

$$\liminf_{n\to\infty} \frac{1}{\log n} \log \mathbf{P}\left(\frac{p^n W_n}{\log n} > t\right) \geqq -\min\{f_i(t) : 0 \leqq i \leqq k\},$$

completing the proof.

Proof of Theorem 2 (a). The assertion will immediately follow from the Corollary applied to the subset $H(0)$, as soon as we show that $W(H(0))$ is the dominant part of W_n.

Since

$$\mathbf{P}\left(\frac{p^n W(H(0))}{\log(a^n)} > t\right) \leqq \mathbf{P}\left(\frac{p^n W_n}{\log(a^n)} > t\right) \leqq \mathbf{P}\left(\frac{p^n W(H(0))}{\log(a^n)} > t\right) + \sum_{i=1}^{n} S_i,$$

where

$$S_i = \sum_{A\in H(i)} \mathbf{P}\left(\frac{p^n T(A)}{\log(a^n)} > t\right) = \sum_{A\in H(i)} \mathbf{P}\left(\frac{T(A)}{A*A} > \frac{\log(a^n)}{p^n A*A} t\right),$$

we just have to estimate S_i, $i \geqq 1$.

Firstly, let $1 \leqq i < mk$. Similarly to (3.3) one can write

$$S_i = \sum_{A\in H(i)} \mathbf{P}\left(\frac{T(A)}{A*A} > t\lambda^i \log(a^n)(1+o(1))\right)$$
$$\leqq 2\binom{n}{i} a^{n-i}(d-a)^i \exp\{-t\lambda^i \log(a^n)(1+o(1))\}$$
$$\leqq \exp\{-(t-1+o(1))\log(a^n)\}.$$

Secondly, let $i \geqq mk$. Since $A*A \leqq p^{n-1}\lambda^{-i}$ holds, by Proposition 3 we have

$$S_i \leqq 2a^n \binom{n}{i} \left(\frac{d-a}{a}\right)^i \exp\{-p\lambda^i t \log(a^n)(1+o(1))\}$$
$$\leqq \binom{n}{i}\left(\frac{d-a}{a}\right)^i \exp\{-(t(p\lambda^i - 1) + (t-1) + o(1))\log(a^n)\}.$$

As we have seen in the lines preceding (3.4), $p\lambda^i - 1 > \frac{i}{k}\lambda^k > \frac{2}{k}i$, thus

$$\sum_{i=mk}^{n} S_i \leqq \exp\{-(t-1+o(1))\log(a^n)\} \sum_{i=mk}^{n} \binom{n}{i} \left(\frac{d-a}{a} a^{-2n/k}\right)^i$$
$$\leqq \left(1 + \frac{d-a}{a} a^{-2n/k}\right)^n \exp\{-(t-1+o(1))\log(a^n)\}.$$

Hence

$$\sum_{i=1}^{n} S_i \leqq \exp\{-(t-1+o(1))\log(a^n)\},$$

and the proof is completed.

ACKNOWLEDGEMENT. This work was supported by the Hungarian National Foundation for Scientific Research Grant No. 1905.

REFERENCES

[1] ALDOUS, D. J., Threshold limits for cover times, *J. Theoret. Probab.* **4** (1991), 197–211. *MR* **91m**:60123

[2] ERDŐS, P. and CHEN, R. W., Random walks on Z_2^n, *J. Multivariate Anal.* **25** (1988), 111–118. *MR* **89d**:60124

[3] LI, S. R., A martingale approach to the study of occurrence of sequence patterns in repeated experiments, *Ann. Probab.* **8** (1980), 1171–1176. *MR* **82c**:60018

[4] MÓRI, T. F., Large deviation results for waiting times in repeated experiments, *Acta Math. Hungar.* **45** (1985), 213–221. *MR* **86h**:60017

[5] MÓRI, T. F., On the waiting time till each of some given patterns occurs as a run, *Probab. Fields* **87** (1991), 313–323. *MR* **91m**:60064

[6] MÓRI, T. F., More on the waiting time till each of some given patterns occurs as a run, *Canad. J. Math.* **42** (1990), 915–932. *MR* **92d**:60035

[7] MÓRI, T. F., Covering with blocks in the non-symmetric case, *J. Theoret. Probab.* **8** (1995), 139–164.

(Received September 7, 1994)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR
VALÓSZÍNŰSÉGELMÉLETI ÉS STATISZTIKAI TANSZÉK
MÚZEUM KRT. 6–8
H–1088 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 237–247

p-VARIATION OF GAUSSIAN PROCESSES WITH STATIONARY INCREMENTS

QI-MAN SHAO

Dedicated to Pál Révész on the occasion of his 60th birthday

Abstract

A general result on p-variation of a mean zero Gaussian process with stationary increments is obtained.

1. Introduction

Let $\{G(t),\, t \geqq 0\}$ be a mean zero Gaussian process with stationary increments. Put

$$\sigma^2(h) = \mathbf{E}(G(x+h) - G(x))^2. \tag{1.1}$$

Let $\pi = \{0 = x_0 < x_1 < \cdots < x_{k_\pi} = a\}$ denote a partition of $[0,a]$ and $m(\pi) = \max_{1 \leqq i \leqq k_\pi} (x_i - x_{i-1})$ denote the length of the largest interval in π.

There has been a great amount of work on the p-variation of G since Lévy [4] initiated the study on the quadratic or 2-variation of the Brownian motion $\{B(t), t \in R^+\}$, that is,

$$\lim_{n\to\infty} \sum_{i=0}^{2^n-1} \left(B\left(\frac{i}{2^n}\right) - B\left(\frac{i+1}{2^n}\right)\right)^2 = 1 \quad \text{a.s.}. \tag{1.2}$$

Dudley [2] generalized (1.2) and proved that for any sequence $\{\pi_n\}$ of interval partitions of $[0,1]$ such that $m(\pi_n) = o(1/\log n)$

$$\lim_{n\to\infty} \sum_{x_i \in \pi_n} (B(x_i) - B(x_{i-1}))^2 = 1 \quad \text{a.s.}, \tag{1.3}$$

while Fernández de la Vega [4] showed that this is no longer true if the condition on $m(\pi_n)$ is relaxed to $m(\pi_n) = O(1/\log n)$. Giné and Klein [3]

1991 *Mathematics Subject Classification*. Primary 60G15, 60G17; Secondary 60F15.
Key words and phrases. Gaussian processes, fractional Brownian motion, p-variation.

The research is partially supported by a National University of Singapore Research Project.

extended (1.3) to Gaussian processes with stationary increments. Recently, Marcus and Rosen [5] considered the p-variation of G and obtained that if $\sigma^2(h)$ is concave on $[0,a]$ and satisfies $\lim_{h\to 0}\sigma(h)/h^{1/p}=c$ for some $p\geqq 2$ and $0<c<\infty$, then for any sequence of partitions $\{\pi_n\}$ of $[0,a]$ such that $m(\pi_n)=o(1/\log^{p/2} n)$

$$\lim_{n\to\infty}\sum_{x_i\in\pi_n}|G(x_i)-G(x_{i-1})|^p=ac^p\mathbf{E}|\eta|^p \quad \text{a.s.} \tag{1.4}$$

and

$$\lim_{n\to\infty}\sum_{x_i\in\pi_n}|G^2(x_i)-G^2(x_{i-1})|^p=\mathbf{E}|\eta|^p(2c)^p\int_0^a|G(x)|^p dx \quad \text{a.s.}. \tag{1.5}$$

Here, and in the sequel, η denotes the standard normal random variable. For further various aspects of p-variation of Gaussian processes as well as its remarkable connection to the local times of symmetric stable processes we refer to Marcus and Rosen [6] and the references therein. The aim of this note is to weaken the condition on $\sigma^2(h)$ and to give a more general and universal result.

THEOREM 1.1. *Let $p>1,\{G(t),t\geqq 0\}$ be a mean zero Gaussian process with stationary increments. Assume that $\sigma^2(h)$ is non-decreasing. Then for any sequence of partitions $\{\pi_n\}$ of $[0,a]$*

$$\lim_{n\to\infty}\sum_{x_i\in\pi_n}\frac{(x_i-x_{i-1})|G(x_i)-G(x_{i-1})|^p}{\sigma^p(x_i-x_{i-1})}=a\,\mathbf{E}|\eta|^p \quad a.s. \tag{1.6}$$

if one of the following two conditions is satisfied:

(A1) *$\sigma^2(h)$ is concave on $[0,a]$, and $m(\pi_n)=o((\log n)^{-1\vee\frac{p}{2}})$;*

(A2) *$\sigma^2(h)$ is convex on $[0,a+\varepsilon_0]$ for some $\varepsilon_0>0$, and*

$$\max_{x_i\in\pi_n}(x_i-x_{i-1})^{\frac{1}{2}+\frac{1}{2}\wedge\frac{1}{p}}/\sigma(x_i-x_{i-1})=o(\log^{-1/2} n).$$

THEOREM 1.2. *Assume that $\sigma^2(h)$ is continuous on $[0,a]$ satisfying*

$$\int_1^\infty\sigma(e^{-z^2})dz<\infty.$$

Then, under the condition of Theorem 1.1

(1.7)

$$\lim_{n\to\infty}\sum_{x_i\in\pi_n}\frac{(x_i-x_{i-1})|G^2(x_i)-G^2(x_{i-1})|^p}{\sigma^p(x_i-x_{i-1})}=a\,2^p\mathbf{E}|\eta|^p\int_0^a|G(x)|^p dx \quad a.s.$$

The following results are immediate consequences of Theorems 1.1 and 1.2.

COROLLARY 1.1. *Assume that* $\sigma^2(h)$ *is non-decreasing and concave on* $[0,a]$ *and satisfies* $\lim_{h\to 0}\sigma(h)/h^\alpha = c$ *for* $0<\alpha\leqq 1/2$ *and* $0<c<\infty$. *Then for* $p>1$ *and for any sequence of partitions* $\{\pi_n\}$ *of* $[0,a]$ *such that* $m(\pi_n) = o((\log n)^{-1\vee\frac{p}{2}})$,

$$\lim_{n\to\infty}\sum_{x_i\in\pi_n}|G(x_i)-G(x_{i-1})|^p(x_i-x_{i-1})^{1-p\alpha} = ac^p\mathbf{E}|\eta|^p \quad a.s. \tag{1.8}$$

and

$$\lim_{n\to\infty}\sum_{x_i\in\pi_n}|G^2(x_i)-G^2(x_{i-1})|^p(x_i-x_{i-1})^{1-p\alpha} = a\mathbf{E}|\eta|^p(2c)^p\int_0^a|G(x)|^p dx \quad a.s.. \tag{1.9}$$

COROLLARY 1.2. *Assume that* $\sigma^2(h)$ *is non-decreasing and convex on* $[0,a+\varepsilon_0]$ *for some* $\varepsilon_0>0$ *and satisfies* $\lim\limits_{h\to 0}\sigma(h)/h^\alpha = c$ *for* $1/2\leqq\alpha<1$ *and* $0<c<\infty$. *Then for* $1<p<2/(2\alpha-1)$ *and for any sequence of partitions* $\{\pi_n\}$ *of* $[0,a]$ *such that* $m(\pi_n) = o\Big((\log n)^{-1/(1-2\alpha+1\wedge 2/p)}\Big)$,

$$\lim_{n\to\infty}\sum_{x_i\in\pi_n}|G(x_i)-G(x_{i-1})|^p(x_i-x_{i-1})^{1-p\alpha} = ac^p\mathbf{E}|\eta|^p \quad a.s. \tag{1.10}$$

$$\lim_{n\to\infty}\sum_{x_i\in\pi_n}|G^2(x_i)-G^2(x_{i-1})|^p(x_i-x_{i-1})^{1-p\alpha} = a\mathbf{E}|\eta|^p(2c)^p\int_0^a|G(x)|^p dx \quad a.s.. \tag{1.11}$$

In particular, for any sequence of partitions $\{\pi_n\}$ *of* $[0,a]$ *such that* $m(\pi_n) = o((\log n)^{-1/(2(1-\alpha))})$,

$$\lim_{n\to\infty}\sum_{x_i\in\pi_n}|G(x_i)-G(x_{i-1})|^{1/\alpha} = ac^{1/\alpha}\mathbf{E}|\eta|^{1/\alpha} \quad a.s. \tag{1.12}$$

$$\lim_{n\to\infty}\sum_{x_i\in\pi_n}|G^2(x_i)-G^2(x_{i-1})|^{1/\alpha} = a\mathbf{E}|\eta|^{1/\alpha}(2c)^{1/\alpha}\int_0^a|G(x)|^{1/\alpha}dx \quad a.s.. \tag{1.13}$$

Applying Corollaries 1.1 and 1.2 to the fractional Brownian motion, we have

COROLLARY 1.3. *Let* $\{G(t), t\geqq 0\}$ *be the fractional Brownian motion of order* α, $0<\alpha<1$, *that is,* $\sigma(h) = h^\alpha$. *Let* $1<p<2/\max\{(2\alpha-1),0\}$ *and*

$r_{p,\alpha} = \max(1, p/2)\, I_{\{0<\alpha\leqq 1/2\}} + 1/(1-2\alpha+1\wedge 2/p)\, I_{\{1/2<\alpha<1\}}$. *Then for any sequence of partitions* $\{\pi_n\}$ *of* $[0,a]$ *such that* $m(\pi_n) = o\left((\log n)^{-r_{p,\alpha}}\right)$,

$$\lim_{n\to\infty} \sum_{x_i\in\pi_n} |G(x_i) - G(x_{i-1})|^p (x_i - x_{i-1})^{1-p\alpha} = a\,\mathbf{E}|\eta|^p \quad a.s.$$

$$\lim_{n\to\infty} \sum_{x_i\in\pi_n} |G^2(x_i) - G^2(x_{i-1})|^p (x_i - x_{i-1})^{1-p\alpha} = a\,2^p\,\mathbf{E}|\eta|^p \int_0^a |G(x)|^p dx \quad a.s..$$

In particular, for any sequence of partitions $\{\pi_n\}$ *of* $[0,a]$ *such that* $m(\pi_n) = o((\log n)^{-1/(2\min(\alpha,1-\alpha))})$,

$$\lim_{n\to\infty} \sum_{x_i\in\pi_n} |G(x_i) - G(x_{i-1})|^{1/\alpha} = a\,\mathbf{E}|\eta|^{1/\alpha} \quad a.s.$$

$$\lim_{n\to\infty} \sum_{x_i\in\pi_n} |G^2(x_i) - G^2(x_{i-1})|^{1/\alpha} = a\,2^{1/\alpha}\,\mathbf{E}|\eta|^{1/\alpha} \int_0^a |G(x)|^{1/\alpha} dx \quad a.s..$$

Corollary 1.1 with $p = 1/\alpha$ coincides Theorem 1.2 of Marcus and Rosen [5]. The results in (1.8) and (1.10) for $p = 2$ overlap Theorem 2 of Giné and Klein [3], but the assumption here looks simpler. We believe results (1.12) and (1.13) are new.

2. Proof

We start with preliminary lemmas, some of which are of independent interest.

Lemma 2.1. *Let* (X, Y) *be a normal random vector with* $\mathbf{E}X = \mathbf{E}Y = 0$. *Then*

$$\mathbf{E}|X|^p|Y|^p \leqq \mathbf{E}|X|^p\mathbf{E}|Y|^p + C_p\,|\mathbf{E}XY|\,\mathbf{E}|X|^{p-1}\,\mathbf{E}|Y|^{p-1}$$

for $p \geqq 1$, *where* $C_p = 2^p(1 + \mathbf{E}|\eta|^{2p}/(\mathbf{E}|\eta|^{p-1})^2)$ *and* η *is the standard normal random variable.*

Proof. Put $\rho = \mathbf{E}(XY)\,/\,\mathbf{E}Y^2$. Using the elementary inequality

$$(1+x)^p \leqq 1 + 2^p(x + x^p) \quad \text{for every } x \geqq 0$$

and noting that $X - \rho Y$ and Y are independent, we have

$$\begin{aligned}
\mathbf{E}|X|^p|Y|^p &= \mathbf{E}|X - \rho Y + \rho Y|^p|Y|^p \\
&\leqq \mathbf{E}\left(|X - \rho Y|^p + 2^p\{|X - \rho Y|^{p-1}|\rho Y| + |\rho Y|^p\}\right)|Y|^p \\
&= \mathbf{E}|X - \rho Y|^p\mathbf{E}|Y|^p + 2^p\{|\rho|\,\mathbf{E}|X - \rho Y|^{p-1}\mathbf{E}|Y|^{p+1} + |\rho|^p\,\mathbf{E}|Y|^{2p}\} \\
&\leqq \mathbf{E}|X|^p\mathbf{E}|Y|^p + 2^p\{|\rho|^p(\mathbf{E}Y^2)^p\mathbf{E}|\eta|^{2p} \\
&\quad + (\mathbf{E}X^2 - \rho^2\mathbf{E}Y^2)^{(p-1)/2}(\mathbf{E}Y^2)^{(p+1)/2}\mathbf{E}|\eta|^{p-1}\mathbf{E}|\eta|^{p+1}\} \\
&\leqq \mathbf{E}|X|^p\mathbf{E}|Y|^p + 2^p\{|\mathbf{E}XY|^p\mathbf{E}|\eta|^{2p} + |\mathbf{E}XY|\,\mathbf{E}|X|^{p-1}\,\mathbf{E}|Y|^{p-1}\} \\
&\leqq \mathbf{E}|X|^p\mathbf{E}|Y|^p + 2^p\{|\mathbf{E}XY|(\mathbf{E}X^2\,\mathbf{E}Y^2)^{(p-1)/2}\mathbf{E}|\eta|^{2p} \\
&\quad + |\mathbf{E}XY|\,\mathbf{E}|X|^{p-1}\,\mathbf{E}|Y|^{p-1}\} \\
&= \mathbf{E}|X|^p\mathbf{E}|Y|^p + 2^p|\mathbf{E}XY|\,\mathbf{E}|X|^{p-1}\,\mathbf{E}|Y|^{p-1}\{1 + \mathbf{E}|\eta|^{2p}/(\mathbf{E}|\eta|^{p-1})^2\}
\end{aligned}$$

as desired.

LEMMA 2.2. *Let $B = (b_{ij})$ be an $n \times n$ real symmetric matrix. Then for any $x = (x_1, \dots, x_n)$*

$$x\,B\,x' \leqq \sum_{i=1}^{n} x_i^2 b_i^*, \tag{2.1}$$

where $b_i^* = \sum_{j=1}^{n} |b_{ij}|$, $i = 1, 2, \dots, n$.

PROOF. Let B^* be a diagonal matrix with diagonal elements $\{b_i^*\}$. It is well-known that $B^* - B$ is a symmetric positive semidefinite matrix. Hence (2.1) follows. □

LEMMA 2.3. *Let $p \geqq 1$, and $\{\xi_i, 1 \leqq i \leqq n\}$ be a Gaussian sequence with mean zero. Then for any $x > 0$ and any sequence of positive numbers $\{a_i, 1 \leqq i \leqq n\}$*

$$\mathbf{P}\left(\left|\left(\sum_{i=1}^{n} a_i|\xi_i|^p\right)^{1/p} - \mathbf{E}\left(\sum_{i=1}^{n} a_i|\xi_i|^p\right)^{1/p}\right| \geqq x\right)$$

$$\leqq \begin{cases} 2\exp\left(-\dfrac{x^2}{2\left(\sum_{i=1}^{n} a_i^{2/(2-p)}\rho_i^{p/(2-p)}\right)^{(2-p)/p}}\right) & \text{if } 1 \leqq p < 2, \\[2ex] 2\exp\left(-\dfrac{x^2}{2\max_{1\leqq i\leqq n}(a_i^{2/p}\rho_i)}\right) & \text{if } p \geqq 2 \end{cases} \tag{2.2}$$

where $\rho_i = \sum_{j=1}^{n} |\mathbf{E}\xi_i\xi_j|$, $i = 1, 2, \dots, n$.

PROOF. Put $q = p/(p-1)$ and define $||b||_q = (\sum_{i=1}^{n} |b_i|^q)^{1/q}$ for $b = (b_1, \dots, b_n)$. It is well-known that

$$\left(\sum_{i=1}^{n} a_i|\xi_i|^p\right)^{1/p} = \sup_{||b||_q \leqq 1} \sum_{i=1}^{n} a_i^{1/p}\, b_i\, \xi_i.$$

By Borel's inequality, we have

$$\begin{aligned}
&\mathbf{P}\left(\left|\left(\sum_{i=1}^n a_i|\xi_i|^p\right)^{1/p} - \mathbf{E}\left(\sum_{i=1}^n a_i|\xi_i|^p\right)^{1/p}\right| \geqq x\right) \\
(2.3)\qquad &= \mathbf{P}\left(\Big| \sup_{\|b\|_q \leqq 1} \sum_{i=1}^n a_i^{1/p} b_i \xi_i - \mathbf{E} \sup_{\|b\|_q \leqq 1} \sum_{i=1}^n a_i^{1/p} b_i \xi_i\Big| \geqq x\right) \\
&\leqq 2\exp\left(-\frac{x^2}{2 \sup_{\|b\|_q \leqq 1} \mathbf{E}(\sum_{i=1}^n a_i^{1/p} b_i \xi_i)^2}\right).
\end{aligned}$$

Let $B = (\rho_{ij})$ be the covariance matrix of $(\xi_1, \ldots, \xi_n)$ and

$$y = (a_1^{1/p} b_1, \ldots, a_n^{1/p} b_n).$$

It follows from Lemma 2.2 that

$$\begin{aligned}
&\sup_{\|b\|_q \leqq 1} \mathbf{E}(\sum_{i=1}^n a_i^{1/p} b_i \xi_i)^2 = \sup_{\|b\|_q \leqq 1} y B y' \\
&\leqq \sup_{\|b\|_q \leqq 1} \sum_{i=1}^n a_i^{2/p} b_i^2 \rho_i \\
(2.4)\qquad &\leqq \begin{cases} \sup_{\|b\|_q \leqq 1} \|b\|_q^2 \left(\sum_{i=1}^n (a_i^{2/p}\rho_i)^{1/(1-2/q)}\right)^{1-2/q} & \text{if } 1 \leqq p < 2, \\ \sup_{\|b\|_q \leqq 1} \max_{i \leqq n}(a_i^{2/p} \rho_i)\|b\|_2^2 & \text{if } \ p \geqq 2 \end{cases} \\
&\leqq \begin{cases} \left(\sum_{i=1}^n (a_i^{2/p}\rho_i)^{1/(1-2/q)}\right)^{1-2/q} & \text{if } 1 \leqq p < 2, \\ \max_{i \leqq n}(a_i^{2/p} \rho_i) & \text{if } \ p \geqq 2. \end{cases}
\end{aligned}$$

This proves (2.2), by (2.3) and (2.4). □

From the Hölder inequality it follows immediately that

$$\mathbf{E}\left(\sum_{i=1}^n a_i|\xi_i|^p\right)^{1/p} \leqq \left(\sum_{i=1}^n a_i \mathbf{E}|\xi_i|^p\right)^{1/p}. \tag{2.5}$$

Our next lemma gives a lower bound of $\mathbf{E}(\sum_{i=1}^n a_i\,|\xi_i|^p)^{1/p}$.

LEMMA 2.4. *Let $p \geqq 1$, $\{\xi_i, 1 \leqq i \leqq n\}$ be a Gaussian sequence with mean zero. Then for any sequence of positive numbers $\{a_i, 1 \leqq i \leqq n\}$*

$$\begin{aligned} &\mathbf{E}\Big(\sum_{i=1}^{n} a_i|\xi_i|^p\Big)^{1/p} \\ (2.6)\qquad &\geqq \left(\sum_{i=1}^{n} a_i\,\mathbf{E}|\xi_i|^p\right)^{1/p}\left(1 + C_p\sum_{i=1}^{n} a_i^2\mathbf{E}|\xi_i|^{2p-2}\rho_i \Big/ \left(\sum_{i=1}^{n} a_i\,\mathbf{E}|\xi_i|^p\right)^2\right)^{(1-p)/p} \end{aligned}$$

where $\rho_i = \sum_{j=1}^{n} |\mathbf{E}\xi_i\,\xi_j|$, and C_p is defined as in Lemma 2.1.

PROOF. By the Lyapunov inequality, for any random variable X

$$(\mathbf{E}|X|^p)^{2p-1} \leqq (\mathbf{E}|X|)^p(\mathbf{E}|X|^{2p})^{(p-1)}.$$

Hence,

$$\begin{aligned} \mathbf{E}\left(\sum_{i=1}^{n} a_i|\xi_i|^p\right)^{1/p} &\geqq \frac{\left(\mathbf{E}\left(\sum_{i=1}^{n} a_i|\xi_i|^p\right)\right)^{(2p-1)/p}}{\left(\mathbf{E}\left(\sum_{i=1}^{n} a_i|\xi_i|^p\right)^2\right)^{(p-1)/p}} \\ (2.7)\qquad &= \frac{\left(\sum_{i=1}^{n} a_i\mathbf{E}|\xi_i|^p\right)^{(2p-1)/p}}{\left(\mathbf{E}\left(\sum_{i=1}^{n} a_i|\xi_i|^p\right)^2\right)^{(p-1)/p}}. \end{aligned}$$

By Lemmas 2.1 and 2.2, we have

$$\begin{aligned} \mathbf{E}\left(\sum_{i=1}^{n} a_i|\xi_i|^p\right)^2 &= \left(\sum_{i=1}^{n} a_i\mathbf{E}|\xi_i|^p\right)^2 + \mathbf{E}\left(\sum_{i=1}^{n} a_i(|\xi_i|^p - \mathbf{E}|\xi_i|^p)\right)^2 \\ &\leqq \left(\sum_{i=1}^{n} a_i\mathbf{E}|\xi_i|^p\right)^2 + C_p\sum_{1\leqq i,j\leqq n} a_i\,a_j\,\mathbf{E}|\xi_i|^{p-1}\,\mathbf{E}|\xi_j|^{p-1}|\mathbf{E}\xi_i\,\xi_j| \\ (2.8)\qquad &\leqq \left(\sum_{i=1}^{n} a_i\mathbf{E}|\xi|^p\right)^2 + C_p\sum_{1\leqq i\leqq n} a_i^2(\mathbf{E}|\xi_i|^{p-1})^2\rho_i \\ &\leqq \left(\sum_{i=1}^{n} a_i\,\mathbf{E}|\xi_i|^p\right)^2 + C_p\sum_{1\leqq i\leqq n} a_i^2\,\mathbf{E}|\xi_i|^{2p-2}\rho_i. \end{aligned}$$

(2.6) now follows from (2.7) and (2.8). □

PROOF OF THEOREM 1.1. We first collect two facts:

(A) If $\sigma^2(h)$ is concave on $[0,a]$, then

$$\mathbf{E}(G(x_2)-G(x_1))(G(x_4)-G(x_3)) \leqq 0 \quad \text{for} \quad 0 \leqq x_1 \leqq x_2 \leqq x_3 \leqq x_4 \leqq a;$$

(B) If $\sigma^2(h)$ is convex on $[0, a+\varepsilon_0]$ for some $\varepsilon_0 > 0$, then

$$\mathbf{E}(G(x_2)-G(x_1))(G(x_4)-G(x_3)) \geqq 0 \quad \text{for} \quad 0 \leqq x_1 \leqq x_2 \leqq x_3 \leqq x_4 \leqq a$$

and there is $K > 0$ such that

$$\sigma^2(x+y) - \sigma^2(x) \leqq Ky \quad \text{for all} \quad 0 \leqq x \leqq x+y \leqq a.$$

Let

$$\mathbf{E}_{\pi_n} = \mathbf{E}\left(\sum_{x_i\in\pi_n} \frac{(x_i - x_{i-1})|G(x_i)-G(x_{i-1})|^p}{\sigma^p(x_i - x_{i-1})}\right)^{1/p},$$

$$\rho_i = \sum_{x_j\in\pi_n} |\mathbf{E}(G(x_i)-G(x_{i-1}))(G(x_j)-G(x_{j-1}))|,$$

$$\delta_n = \begin{cases} \left(\sum_{x_i\in\pi_n} \left(\frac{x_i-x_{i-1}}{\sigma^p(x_i-x_{i-1})}\right)^{2/(2-p)} \rho_i^{p/(2-p)}\right)^{(2-p)/p} & \text{for} \quad 1 \leqq p < 2, \\ \max_{x_i\in\pi_n} \rho_i \left(\frac{x_i-x_{i-1}}{\sigma^p(x_i-x_{i-1})}\right)^{2/p} & \text{for} \quad p > 2. \end{cases}$$

We below prove that

$$\delta_n = o(1/\log n) \quad \text{as} \quad n \to \infty. \tag{2.9}$$

The proof is formulated into two cases.

Case 1. σ^2 is concave on $[0,a]$. From (A) we get

$$\begin{aligned} \rho_i =& \sigma^2(x_i - x_{i-1}) - \sum_{j\neq i, x_j\in\pi_n} \mathbf{E}(G(x_i)-G(x_{i-1}))(G(x_j)-G(x_{j-1})) \\ =& \sigma^2(x_i - x_{i-1}) - \mathbf{E}(G(x_i)-G(x_{i-1}))(G(a)-G(x_i)) \\ & - \mathbf{E}(G(x_i)-G(x_{i-1}))(G(x_{i-1})-G(0)) \\ \leqq& 2\sigma^2(x_i - x_{i-1}). \end{aligned} \tag{2.10}$$

Thus, by (A1)

(2.11)

$$\delta_n \leq \begin{cases} \left(\sum_{x_i \in \pi_n} \left(\frac{x_i - x_{i-1}}{\sigma^p(x_i - x_{i-1})}\right)^{2/(2-p)} (2\sigma^2(x_i - x_{i-1}))^{p/(2-p)} \right)^{(2-p)/p} & \text{for } 1 \leqq p < 2 \\ \max_{x_i \in \pi_n} 2\sigma^2(x_i - x_{i-1}) \left(\frac{x_i - x_{i-1}}{\sigma^p(x_i - x_{i-1})}\right)^{2/p} & \text{for } p > 2 \end{cases}$$

$$\leqq \begin{cases} 2\left(\sum_{x_i \in \pi_n} (x_i - x_{i-1})^{2/(2-p)} \right)^{(2-p)/p} & \text{for } 1 \leqq p < 2 \\ 2 \max_{x_i \in \pi_n} (x_i - x_{i-1})^{2/p} & \text{for } p > 2 \end{cases}$$

$$\leqq \begin{cases} 2m(\pi_n) \left(\sum_{x_i \in \pi_n} (x_i - x_{i-1}) \right)^{(2-p)/p} & \text{for } 1 \leqq p < 2 \\ 2m(\pi_n)^{2/p} & \text{for } p > 2 \end{cases}$$

$$= o(1/\log n).$$

This proves (2.9).

Case 2. σ^2 is convex on $[0, a + \varepsilon_0]$. Using (B), we have

$$\begin{aligned} \rho_i &= \mathbf{E}(G(x_i) - G(x_{i-1})) \sum_{x_j \in \pi_n} (G(x_j) - G(x_{j-1})) \\ &= \mathbf{E}(G(x_i) - G(x_{i-1}))(G(a) - G(0)) \\ &= \frac{1}{2}(\sigma^2(a - x_{i-1}) - \sigma^2(a - x_i) + \sigma^2(x_i) - \sigma^2(x_{i-1})) \\ &\leqq K(x_i - x_{i-1}). \end{aligned} \tag{2.12}$$

Similar to (2.11), by (A2)

(2.13)

$$\delta_n \leq \begin{cases} \left(\sum_{x_i \in \pi_n} \left(\frac{x_i - x_{i-1}}{\sigma^p(x_i - x_{i-1})}\right)^{2/(2-p)} (K(x_i - x_{i-1}))^{p/(2-p)} \right)^{(2-p)/p} & \text{for } 1 \leqq p < 2 \\ K \max_{x_i \in \pi_n} (x_i - x_{i-1}) \left(\frac{x_i - x_{i-1}}{\sigma^p(x_i - x_{i-1})}\right)^{2/p} & \text{for } p > 2 \end{cases}$$

$$\leqq \begin{cases} K \max_{x_j \in \pi_n} \frac{(x_j - x_{j-1})^2}{\sigma^2(x_j - x_{j-1})} \left(\sum_{x_i \in \pi_n} (x_i - x_{i-1}) \right)^{2/(2-p)} & \text{for } 1 \leqq p < 2 \\ K \max_{x_i \in \pi_n} \frac{(x_i - x_{i-1})^{1+2/p}}{\sigma^2(x_i - x_{i-1})} & \text{for } p > 2 \end{cases}$$

$$= o(1/\log n),$$

which follows (2.9), too.

From Lemma 2.3 and (2.9) we obtain for any $\varepsilon > 0$

$$\begin{aligned}(2.14)\qquad &\mathbf{P}\left(\left|\left(\sum_{x_i\in\pi_n}\frac{(x_i-x_{i-1})|G(x_i)-G(x_{i-1})|^p}{\sigma^p(x_i-x_{i-1})}\right)^{1/p}-\mathbf{E}_{\pi_n}\right|>\varepsilon\right)\\ &\qquad \leqq 2\exp\left(-\frac{\varepsilon^2}{2\delta_n}\right)=2\exp\left(-\frac{\varepsilon^2\log n}{o(1)}\right).\end{aligned}$$

Therefore, by the Borel–Cantelli lemma

$$(2.15)\qquad \lim_{n\to\infty}\left(\sum_{x_i\in\pi_n}\frac{(x_i-x_{i-1})|G(x_i)-G(x_{i-1})|^p}{\sigma^p(x_i-x_{i-1})}\right)^{1/p}-\mathbf{E}_{\pi_n}=0\quad \text{a.s..}$$

Since

$$\sum_{x_i\in\pi_n}\frac{(x_i-x_{i-1})\mathbf{E}|G(x_i)-G(x_{i-1})|^p}{\sigma^p(x_i-x_{i-1})}=a\,\mathbf{E}|\eta|^p,$$

to finish the proof, by (2.14) and (2.5), it suffices to show that

$$(2.16)\qquad \liminf_{n\to\infty}\mathbf{E}_{\pi_n}\geqq(a\,\mathbf{E}|\eta|^p)^{1/p}.$$

In terms of (2.10) and (2.12), we have

$$\begin{aligned}&\sum_{x_i\in\pi_n}\frac{(x_i-x_{i-1})^2\mathbf{E}|G(x_i)-G(x_{i-1})|^{2p-2}}{\sigma^{2p}(x_i-x_{i-1})}\rho_i\\ &=\mathbf{E}|\eta|^{2p-2}\sum_{x_i\in\pi_n}\frac{(x_i-x_{i-1})^2\rho_i}{\sigma^2(x_i-x_{i-1})}\\ (2.17)\qquad &\leqq\begin{cases}2\mathbf{E}|\eta|^{2p-2}\sum_{x_i\in\pi_n}(x_i-x_{i-1})^2 & \text{if } \sigma^2 \text{ is concave}\\ 2\,K\,\mathbf{E}|\eta|^{2p-2}\sum_{x_i\in\pi_n}\frac{(x_i-x_{i-1})^3}{\sigma^2(x_i-x_{i-1})} & \text{if } \sigma^2 \text{ is convex}\end{cases}\\ &\leqq\begin{cases}2\,a\,\mathbf{E}|\eta|^{2p-2}m(\pi_n) & \text{if } \sigma^2 \text{ is concave}\\ 2\,a\,K\,\mathbf{E}|\eta|^{2p-2}\max_{x_i\in\pi_n}\frac{(x_i-x_{i-1})^2}{\sigma^2(x_i-x_{i-1})} & \text{if } \sigma^2 \text{ is convex}\end{cases}\\ &=o(1)\quad\text{as}\quad n\to\infty.\end{aligned}$$

Hence, (2.15) holds, by Lemma 2.4 and (2.16).

The proof is now complete. □

PROOF OF THEOREM 1.2. The proof follows from that of Theorem 1.2 of Marcus and Rosen [5].

REFERENCES

[1] Fernández de la Vega, W., On almost sure convergence of quadratic Brownian variation, *Ann. Probab.* **2** (1974), 551–552. *MR* **50** #11484

[2] Dudley, R. M., Sample functions of the Gaussian process, *Ann. Probab.* **1** (1973), 66–103. *MR* **49** #11605

[3] Giné, E. and Klein, R., On quadratic variation of processes with Gaussian increments, *Ann. Probab.* **3** (1975), 716–721. *MR* **51** #14239

[4] Lévy, P., Le mouvement brownien plan, *Amer. J. Math.* **62** (1940), 487–550. *MR* **2**–107

[5] Marcus, M. B. and Rosen, J., p-variation of the local times of symmetric stable processes and of Gaussian processes with stationary increments, *Ann. Probab.* **20** (1992), 1685–1713. *MR* **94c**:60124

[6] Marcus, M. B. and Rosen, J., Sample path properties of the local times of strongly symmetric Markov processes via Gaussian processes, *Ann. Probab.* **20** (1992), 1603–1684. *MR* **93k**:60089

(Received December 21, 1994)

DEPARTMENT OF MATHEMATICS
NATIONAL UNIVERSITY OF SINGAPORE
SINGAPORE 0511
REPUBLIC OF SINGAPORE
matsqm@leonis.nus.sg

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 249–297

AN ELEMENTARY INTRODUCTION TO THE WIENER PROCESS AND STOCHASTIC INTEGRALS

T. SZABADOS

Dedicated to P. Révész for his sixtieth birthday

Abstract

An elementary construction of the Wiener process is discussed, based on a proper sequence of simple symmetric random walks that uniformly converge on bounded intervals, with probability 1. This method is a simplification of F.B. Knight's and P. Révész's. The same sequence is applied to give elementary (Lebesgue-type) definitions of Itô and Stratonovich sense stochastic integrals and to prove the basic Itô formula. The resulting approximating sums converge with probability 1. As a by-product, new elementary proofs are given for some properties of the Wiener process, like the almost sure non-differentiability of the sample-functions. The purpose of using elementary methods almost exclusively is twofold: first, to provide an introduction to these topics for a wide audience; second, to create an approach well-suited for generalization and for attacking otherwise hard problems.

1. Introduction

The *Wiener process* is undoubtedly one of the most important stochastic processes, both in the theory and in the applications. Originally it was introduced as a mathematical model of *Brownian motion*, a random zigzag motion of microscopic particles suspended in liquid, discovered by the English botanist Brown in 1826. An amazing number of first class scientists like Bachelier, Einstein, Smoluchowski, Wiener, and Lévy, to mention just a few, contributed to the theory of Brownian motion. In the course of the evolution of probability theory it became clear that the Wiener process is a basic tool for many limit theorems and also a natural model of many phenomena involving randomness, like noise, random fluctuations or perturbations.

The Wiener process is a natural model of Brownian motion. It describes a random, but continuous motion of a particle, subjected to the influence of a large number of chaotically moving molecules of the liquid. Any displacement of the particle over an interval of time as a sum of many almost independent small influences is normally distributed with expectation zero and variance proportional to the length of the time interval. Displacements over disjoint time intervals are independent.

1991 *Mathematics Subject Classification*. Primary 60J65, 60H05; Secondary 60J15, 60F15.

Key words and phrases. Brownian motion, Wiener process, random walks, stochastic integrals, Itô formula.

0081-6906/96/$ 5.00

The most basic types of *stochastic integrals* were introduced by K. Itô and R. L. Stratonovich as tools for investigating stochastic differential equations, that is, differential equations containing random functions. Not surprisingly, the Wiener process is one of the corner stones the theory of stochastic integrals and differential equations was built on.

Stochastic differential equations are applied under similar conditions as differential equations in general. The advantage of the stochastic model is that it can accommodate noise or other randomly changing input and effects, which is a necessity in many applications. When solving a stochastic differential equation one has to integrate a function with respect to the increments of a stochastic process like the Wiener process. In such a case the classical methods of integration cannot be applied directly because of the "strange" behaviour of the increments of the Wiener and similar processes.

A main purpose of this paper is to provide an elementary introduction to the aforementioned topics. The discussion of the Wiener process is based on a nice, natural construction of P. Révész [6, Section 6.2], which is essentially a simplified version of F.B. Knight's [4, Section 1.3]. We use a proper sequence of simple random walks that converge to the Wiener process. Then an elementary definition and discussion of stochastic integrals is given, based on [8], which uses the same sequence of random walks.

The level of the paper is (hopefully) available to any good student who has taken a usual calculus sequence and an introductory course in probability. Our general reference will be W. Feller's excellent, elementary textbook [2]. Anything that goes beyond the material of that book will be discussed here in detail. I would like to convince the reader that these important and widely used topics are natural and feasible supplements to a strong introductory course in probability; this way a much wider audience could get acquainted with them. However, I have to warn the non-expert reader that "elementary" is not a synonym of "easy" or "short".

To encourage the reader it seems worthwhile to emphasize a very useful feature of elementary approaches: in many cases, elementary methods are easier to generalize or to attack otherwise hard problems.

2. Random walks

The simplest (and crudest) model of Brownian motion is a *simple symmetric random walk* in one dimension, hereafter *random walk* for brevity.

A particle starts from the origin and steps one unit either to the left or to the right with equal probabilities 1/2, in each unit of time. Mathematically, we have a sequence $X_1, X_2, \ldots$ of independent and identically distributed random variables with

$$\mathbf{P}\{X_n = 1\} = \mathbf{P}\{X_n = -1\} = 1/2 \quad (n = 1, 2, \ldots),$$

and the position of the particle at time n (that is, the random walk) is given by the partial sums

$$S_0 = 0, \; S_n = X_1 + X_2 + \cdots + X_n \quad (n = 1, 2, \dots). \tag{1}$$

The notation $X(n)$ and $S(n)$ will be used instead of X_n and S_n where it seems to be advantageous.

A bit of terminology: a *stochastic process* is a collection $Z(t)$ $(t \in T)$ of random variables defined on a *sample space* Ω. Usually T is a subset of the real line and t is called "time". An important concept is that of a *sample-function*, that is, a randomly selected path of a stochastic process. A sample-function of a stochastic process $Z(t)$ can be denoted by $Z(t;\omega)$, where $\omega \in \Omega$ is fixed, but the "time" t is not.

To visualize the graph of a sample-function of the random walk one can use a broken line connecting the vertices (n, S_n), $n = 1, 2, \dots$ (Figure 1). This way the sample-functions are extended from the set of the non-negative integers to continuous functions on the interval $[0, \infty)$:

$$S(t) = S(n) + (t - n)X_{n+1} \; (n \leqq t < n + 1; \; n = 0, 1, 2, \dots). \tag{2}$$

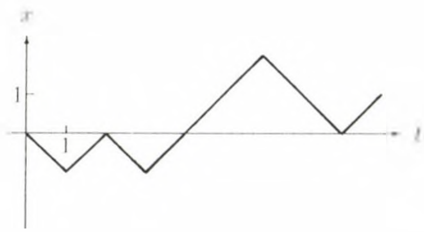

Fig. 1. The graph of a sample-function of $S(t)$.

It is easy to evaluate the expectation and variance of S_n:

$$\mathbf{E}(S_n) = \sum_{k=1}^{n} \mathbf{E}(X_k) = 0, \qquad \mathbf{Var}(S_n) = \sum_{k=1}^{n} \mathbf{E}(X_k^2) = n. \tag{3}$$

The distribution of S_n is a linearly transformed symmetric binomial distribution [2, Section III,2]. Each path (broken line) of length n has probability $1/2^n$. The number of paths going to the point (n, r) from the origin is equal to the number of choosing $(n + r)/2$ steps to the right out of n steps. Consequently,

$$\mathbf{P}\left\{S_n = r\right\} = \binom{n}{(n+r)/2} \frac{1}{2^n} \qquad (|r| \leqq n).$$

The binomial coefficient here is considered to be zero when $n+r$ is not divisible by 2. Equivalently, $S_n = 2B_n - n$, where B_n is a symmetric ($p=1/2$) binomial random variable, $\mathbf{P}\{B_n = k\} = \binom{n}{k}2^{-n}$.

An elementary computation shows that for n large, the binomial distribution can be approximated by the normal distribution, see [2, Section VII,2]. What is shown there that for even numbers $n = 2\nu$ and $r = 2k$, if $n \to \infty$ and $|r| < K_n = o(n^{2/3})$, one has

$$\text{(4)} \quad \mathbf{P}\{S_n = r\} = \binom{n}{(n+r)/2}\frac{1}{2^n} = \binom{2\nu}{\nu+k}\frac{1}{2^{2\nu}} \sim \frac{1}{\sqrt{\pi\nu}}e^{-k^2/\nu} = 2h\phi(rh),$$

where $h = 1/\sqrt{n}$ and $\phi(x) = (1/\sqrt{2\pi})e^{-x^2/2}$ $(-\infty < x < \infty)$, the standard normal density function. Note that for odd numbers $n = 2\nu+1$ and $r = 2k+1$ (4) can be proved similarly as for even numbers.

Here and later we adopt the usual notations $a_n \sim b_n$ for $\lim_{n\to\infty} a_n/b_n = 1$ (a_n and b_n are *asymptotically equal*), and $a_n = o(b_n)$ for $\lim_{n\to\infty} a_n/b_n = 0$.

Equation (4) easily implies a special case of the *central limit theorem* and of the *large deviation theorem* [2, Sections VII,3 and 6]:

THEOREM 1. (a) *For any real x fixed and $n \to \infty$ we have*

$$\mathbf{P}\left\{S_n/\sqrt{n} \leqq x\right\} \to \Phi(x),$$

where $\Phi(x) = (1/\sqrt{2\pi})\int_{-\infty}^{x} e^{-u^2/2}\,du$ $(-\infty < x < \infty)$ is the standard normal distribution function.

(b) *If $n \to \infty$ and $x_n \to \infty$ so that $x_n = o(n^{1/6})$, then*

$$\mathbf{P}\left\{S_n/\sqrt{n} \geqq x_n\right\} \sim 1 - \Phi(x_n),$$
$$\mathbf{P}\left\{S_n/\sqrt{n} \leqq -x_n\right\} \sim \Phi(-x_n) = 1 - \Phi(x_n). \qquad \square$$

For us the most essential statement of the theorem is that when x_n goes to infinity (slower than $n^{1/6}$), then the two sides of (6) tend to zero equally fast, in fact very fast. For, to estimate $1 - \Phi(x)$ for x large, one can use the following inequality, see [2, Section VII,1],

$$\text{(5)} \qquad 1 - \Phi(x) < \frac{1}{x\sqrt{2\pi}}e^{-x^2/2} \qquad (x > 0).$$

Thus fixing an $\epsilon > 0$, say $\epsilon = 1/2$, there exists an integer $n_0 > 0$ such that

$$\text{(6)} \qquad \mathbf{P}\left\{\left|\frac{S_n}{\sqrt{n}}\right| \geqq x_n\right\} \leqq \frac{2(1+\epsilon)}{x_n\sqrt{2\pi}}e^{-x_n^2/2} \leqq e^{-x_n^2/2},$$

for $n \geq n_0$, whenever $x_n \to \infty$ and $x_n = o(n^{1/6})$ as $n \to \infty$. It is important to observe that though S_n can take on every integer from $-n$ to n with positive probability, the event $\{|S_n| > x_n\sqrt{n}\}$ is negligible as $n \to \infty$.

But what can we do if n does not go to ∞, or if the condition $x_n = o(n^{1/6})$ does not hold? Then a simple, but still powerful tool, *Chebyshev's inequality* can be used. A standard form of Chebyshev's inequality [2, Section IX,6] is

$$\mathbf{P}\{|X - \mathbf{E}(X)| \geqq t\} \leqq \frac{\mathbf{Var}(X)}{t^2},$$

for any $t > 0$, supposing $\mathbf{Var}(X)$ is finite. An other form that can be proved similarly is

$$\mathbf{P}\{|X| \geqq t\} \leqq \frac{\mathbf{E}(|X|)}{t}, \tag{7}$$

for any $t > 0$ if $\mathbf{E}(X)$ is finite. If the *kth moment* of X, $\mathbf{E}(X^k)$ is finite ($k > 0$), then one can apply (7) to $|X|^k$ getting

$$\mathbf{P}\{|X| \geqq t\} = \mathbf{P}\left\{|X|^k \geqq t^k\right\} \leqq \frac{\mathbf{E}(|X|^k)}{t^k},$$

for any $t > 0$.

One can even get an upper bound going to 0 exponentially fast as $t \to \infty$ if $\mathbf{E}(e^{uX})$, the *moment generating function of* X, is finite for some $u_0 > 0$. For then, by (7),

$$\mathbf{P}\{X \geqq t\} = \mathbf{P}\{u_0 X \geqq u_0 t\} = \mathbf{P}\left\{e^{u_0 X} \geqq e^{u_0 t}\right\} \leqq e^{-u_0 t}\mathbf{E}(e^{u_0 X}), \tag{8}$$

for any $t > 0$.

Analogously, if $\mathbf{E}(e^{-u_0 X})$ is finite for some $u_0 > 0$, then

$$\mathbf{P}\{X \leqq -t\} = \mathbf{P}\{-u_0 X \geqq u_0 t\} = \mathbf{P}\left\{e^{-u_0 X} \geqq e^{u_0 t}\right\} \leqq e^{-u_0 t}\mathbf{E}(e^{-u_0 X}), \tag{9}$$

for any $t > 0$. Combining (8) and (9), one gets

$$\mathbf{P}\{|X| \geqq t\} = \mathbf{P}\{X \geqq t\} + \mathbf{P}\{X \leqq -t\} \leqq e^{-u_0 t}\left(\mathbf{E}(e^{u_0 X}) + \mathbf{E}(e^{-u_0 X})\right), \tag{10}$$

for any $t > 0$ if the moment generating function is finite both at u_0 and at $-u_0$.

Now, it is easy to find the moment generating function of one step of the random walk:

$$\mathbf{E}(e^{uX_k}) = e^u(1/2) + e^{-u}(1/2) = \cosh u.$$

Hence, using the independence of the steps, one obtains the moment generating function of the random walk S_n as

$$\mathbf{E}(e^{uS_n}) = \mathbf{E}(e^{u\sum_{k=1}^n X_k}) = \mathbf{E}(\prod_{k=1}^n e^{uX_k}) = (\cosh u)^n \qquad (-\infty < u < \infty,\ n \geqq 0). \tag{11}$$

Since $\cosh u$ is an even function and $\cosh 1 < 2$, (10) implies that

$$\mathbf{P}\{|S_n| \geqq t\} \leqq 2 \cdot 2^n e^{-t} \qquad (t > 0,\ n \geqq 0). \tag{12}$$

3. Waiting times

In the sequel we need the distribution of the random time τ when a random walk first hits either the point $x=2$ or -2:

$$\tau=\tau_1=\min\{n:|S_n|=2\}. \tag{13}$$

To find the probability distribution of τ, imagine the random walk as a sequence of pairs of steps. These (independent) pairs can be classified either as a "return": $(1,-1)$ or $(-1,1)$, or as a "change of magnitude 2": $(1,1)$ or $(-1,-1)$. Both cases have the same probability $1/2$.

Clearly, it has zero probability that τ is equal to an odd number. The event $\{\tau=2j\}$ occurs exactly when $j-1$ "returns" are followed by a "change of magnitude 2". Because of the independence of the pairs of steps, $\mathbf{P}\{\tau=2j\}=1/2^j$. It means that $\tau=2Y$, where Y has geometric distribution with parameter $p=1/2$,

$$\mathbf{P}\{\tau=2j\}=\mathbf{P}\{Y=j\}=1/2^j \qquad (j\geqq 1). \tag{14}$$

Hence,

$$\mathbf{E}(\tau)=2\mathbf{E}(Y)=2(1/p)=4,\ \mathbf{Var}(\tau)=2^2\mathbf{Var}(Y)=2^2(1-p)/p^2=8. \tag{15}$$

An important consequence is that with probability 1, a random walk sooner or later hits 2 or -2:

$$\mathbf{P}\{\tau<\infty\}=\sum_{j=1}^{\infty}(1/2^j)=1.$$

It is also quite obvious that

$$\mathbf{P}\{S(\tau)=2\}=\mathbf{P}\{S(\tau)=-2\}=1/2. \tag{16}$$

This follows from the symmetry of the random walk. If we reflect $S(t)$ to the time axis, the resulting process $S^*(t)$ is also a random walk. Its corresponding τ^* is equal to τ, and the event $\{S^*(\tau)=2\}$ is the same as $\{S(\tau)=-2\}$. Since $S^*(t)$ is just the same sort of random walk as $S(t)$, we have $\mathbf{P}\{S^*(\tau)=2\}=\mathbf{P}\{S(\tau)=2\}$ as well.

Another way to show (16) is to use the fact that the waiting time τ has countable many possible values and for any specific value we have symmetry:

$$\begin{aligned}\mathbf{P}\{S(\tau)=2\}&=\sum_{j=1}^{\infty}\mathbf{P}\{S(2j)=2\mid\tau=2j\}\,\mathbf{P}\{\tau=2j\}\\&=\sum_{j=1}^{\infty}\mathbf{P}\{A_{2j-2},X_{2j}=X_{2j-1}=1\mid A_{2j-2},X_{2j}=X_{2j-1}\}\,\mathbf{P}\{\tau=2j\}\\&=(1/2)\sum_{j=1}^{\infty}\mathbf{P}\{\tau=2j\}=1/2,\end{aligned}$$

where A_{2j-2} denotes the event that each of the first $j-1$ pairs is a "return", i.e. $A_{2j-2} = \{X_2 = -X_1, \dots, X_{2j-2} = -X_{2j-3}\}$, $A_0 = \emptyset$.

We mention that (16) illustrates a consequence of the so-called optional sampling theorem, too: $\mathbf{E}(S(\tau)) = 2\mathbf{P}\{S(\tau) = 2\} + (-2)\mathbf{P}\{S(\tau) = -2\} = 0$, which is the same as the expectation of $S(t)$.

We also need the probability of the event that a random walk starting from the point $x = 1$ hits $x = 2$ before hitting $x = -2$. This is equal to the conditional probability $\mathbf{P}\{S(\tau) = 2 \mid X_1 = 1\}$. If $X_1 = 1$, then $X_2 = 1$ with probability $1/2$, and then $\tau = 2$ and $S(\tau) = 2$ as well: $\mathbf{P}\{S(\tau) = 2, \tau = 2 \mid X_1 = 1\} = = 1/2$.

On the other hand, if $X_1 = 1$, then $\tau > 2$ if and only if $X_2 = -1$, with probability $1/2$. that is, at the second step the walk returns the origin and starts "from scratch". Then by (16), it has probability $1/2$ that the random walk hits 2 sooner than -2: $\mathbf{P}\{S(\tau) = 2, \tau > 2 \mid X_1 = 1\} = 1/4$. Therefore

$$\begin{aligned} &\mathbf{P}\{S(\tau) = 2 \mid X_1 = 1\} \\ (17) \qquad &= \mathbf{P}\{S(\tau) = 2, \tau = 2 \mid X_1 = 1\} + \mathbf{P}\{S(\tau) = 2, \tau > 2 \mid X_1 = 1\} \\ &= (1/2) + (1/4) = 3/4. \end{aligned}$$

It also follows then that

$$\mathbf{P}\{S(\tau) = -2 \mid X_1 = 1\} = 1 - (3/4) = 1/4. \tag{18}$$

(16), (17), and (18) are special cases of ruin probabilities [2, Section XIV,2]. For example, it can be shown that the probability that a random walk hits the level $a > 0$ before hitting the level $-b < 0$ is $b/(a+b)$.

Extending definition (13) of τ, for $k = 1, 2, \dots$ we recursively define

$$\tau_{k+1} = \min\{n : n > 0, |S(T_k + n) - S(T_k)| = 2\},$$

where

$$T_k = T(k) = \tau_1 + \tau_2 + \cdots + \tau_k. \tag{19}$$

Then each τ_k has the same distribution as $\tau = \tau_1$. For,

$$\begin{aligned} &\mathbf{P}\{\tau_{k+1} = 2j \mid T_k = 2m\} \\ &\quad = \mathbf{P}\{\min\{n : n > 0, |S(2m+n) - S(2m)| = 2\} = 2j \mid T_k = 2m\} \\ &\quad = \mathbf{P}\{\min\{n : n > 0, |S(n)| = 2\} = 2j\} = \mathbf{P}\{\tau_1 = 2j\} = 1/2^j, \end{aligned}$$

where $k \geqq 1$, $j \geqq 1$, and $m \geqq 1$ are arbitrary. The second equality above follows from two facts. First, each increment $S(2m+n) - S(2m)$ is independent of the event $\{T_k = 2m\}$, because the increment depends only on the random variables X_i $(2m+1 \leqq i \leqq 2m+n)$, while the event $\{T_k = 2m\}$ is determined exclusively by the random variables X_i $(1 \leqq i \leqq 2m)$, the corresponding "past". Second, each increment $S(2m+n) - S(2m)$ has the same distribution as $S(n)$, since both of them is a sum of n independent X_i. Hence, τ_{k+1} is independent of T_k (and also of any $\tau_i, i \leqq k$), so indeed, $\mathbf{P}\{\tau_{k+1} = 2j\} = 1/2^j$ $(j \geqq 1)$.

We also need the distribution of the random time T_k required by k changes of magnitude 2 along the random walk. In other words, $S(t)$ hits even integers (different from the previous one) exclusively at the time instants $T_1, T_2, \ldots$. To find the probability distribution of T_k, imagine the random walk again as a sequence of independent pairs of steps, "returns" and "changes of magnitude 2", both types having probability 1/2. The number of cases the event $\{T_k = 2j\}$ $(j \geqq k)$ can occur is equal to the number of choices of $k-1$ pairs out of $j-1$ where a change of magnitude 2 occurs, before the last pair, which is necessarily a change of magnitude 2. Therefore

$$\mathbf{P}\{T_k = 2j\} = \binom{j-1}{k-1} \frac{1}{2^j} \qquad (j \geqq k \geqq 1). \tag{20}$$

It means that $T_k = 2N_k$, where N_k has a negative binomial distribution with $p = 1/2$ [2, Section VI,8].

All this also follows from the fact that $N_k = T_k/2$ is the sum of k independent, geometrically distributed random variables with parameter $p = 1/2$, see (14) and (19): $N_k = Y_1 + Y_2 + \cdots + Y_k$ $(Y_j = \tau_j/2)$. Then T_k is finite valued with probability 1 and the expectation and variance of T_k easily follows from (15) and (19):

$$\mathbf{E}(T_k) = k\mathbf{E}(\tau) = 4k, \ \mathbf{Var}(T_k) = k\mathbf{Var}(\tau) = 8k. \tag{21}$$

It is worth mentioning that T_k is a *stopping time* for each $k \geqq 1$. By definition, it means that any event of the form $\{T_k \leqq j\}$ depends exclusively on the corresponding "past" $S(t)$ $(t \leqq j)$. In other words, $S_1, \ldots, S_j$ determine whether $\{T_k \leqq j\}$ occurs or not.

Fortunately, the central limit and the large deviation theorems (see Theorem 1) can be proved for negative binomial distributions in the same fashion as for binomial distributions.

THEOREM 2. (a) *For any real x fixed and $k \to \infty$ we have*

$$\mathbf{P}\left\{\frac{T_k - 4k}{\sqrt{8k}} \leqq x\right\} \to \Phi(x).$$

(b) *If $k \to \infty$ and $x_k \to \infty$ so that $x_k = o(k^{1/6})$, then*

$$\mathbf{P}\left\{\frac{T_k - 4k}{\sqrt{8k}} \geqq x_k\right\} \sim 1 - \Phi(x_k),$$

$$\mathbf{P}\left\{\frac{T_k - 4k}{\sqrt{8k}} \leqq -x_k\right\} \sim \Phi(-x_k) = 1 - \Phi(x_k).$$

PROOF. The normal approximation (4) is applicable to negative binomial distributions, too: if $r = 2j$ and $k \to \infty$, then

$$\mathbf{P}\{T_k = r\} = \binom{j-1}{k-1}\frac{1}{2^j} = \frac{1}{2}\binom{j-1}{\frac{(j-1)+(2k-j-1)}{2}}\frac{1}{2^{j-1}}$$
$$\sim \frac{1}{2}\frac{1}{\sqrt{\pi(j-1)/2}}\exp\left(-\frac{\left(\frac{2k-j-1}{2}\right)^2}{(j-1)/2}\right) \tag{22}$$
$$= \frac{1}{\sqrt{\pi(r-2)}}\exp\left(-\frac{(r-4k+2)^2}{4r-8}\right),$$

supposing $|2k-j-1| = o((j-1)^{2/3})$, or equivalently,

$$|r-4k| = o(k^{2/3}). \tag{23}$$

A routine computation shows that (22) is asymptotically equal to

$$\sim \frac{1}{\sqrt{4k\pi}}\exp\left(-\frac{(r-4k)^2}{16k}\right),$$

when $k \to \infty$ and (23) holds. Therefore we get an analogue of (4): if $k \to \infty$ and r is any even number such that $|r-4k| < K_k = o(k^{2/3})$,

$$\mathbf{P}\{T_k = r\} \sim 2h\phi((r-4k)h), \qquad h = 1/\sqrt{8k}, \tag{24}$$

where ϕ denotes the standard normal density function.

Then in the same way as the statements of Theorem 1 are obtained from (4) in [2, Sections VII,3 and 6] one can get the present theorem from (24). Here we recall only the basic step of the argument:

$$\mathbf{P}\left\{x_1 \leqq \frac{T_k - 4k}{\sqrt{8k}} \leqq x_2\right\} \sim \sum_{\{r:\ x_1 \leqq (r-4k)h \leqq x_2,\ r \text{ is even}\}} 2h\phi((r-4k)h)$$
$$\to \int_{x_1}^{x_2}\phi(t)\,dt = \Phi(x_2) - \Phi(x_1),$$

for any x_1, x_2, when $k \to \infty$ and so $h \to 0$. The simple meaning of this is that Riemann sums converge to the corresponding integral. □

In the same manner as the large deviation inequality (6) was obtained for S_n, Theorem 2(b) and (5) imply a large deviation type inequality for T_k:

$$\mathbf{P}\left\{\left|\frac{T_k - 4k}{\sqrt{8k}}\right| \geqq x_k\right\} \leqq e^{-x_k^2/2}, \tag{25}$$

for $k \geqq k_0$, supposing $x_k \to \infty$ and $x_k = o(k^{1/6})$ as $k \to \infty$.

Like in case of S_n, with T_k too we need a substitute for the large deviation inequality if the assumptions $k \to \infty$ or $x_k = o(k^{1/6})$ do not hold. The moment generating function of τ_n is simple:

$$\mathbf{E}(e^{u\tau_n}) = \sum_{j=1}^{\infty} e^{u2j} \frac{1}{2^j} = \frac{e^{2u}/2}{1-(e^{2u}/2)} = \frac{1}{2e^{-2u}-1}. \tag{26}$$

This function is finite if $u < \log\sqrt{2}$. Here and afterwards log denotes logarithm with base e.

Now the moment generating function of T_k follows from the independence of the τ_n's as

$$\mathbf{E}(e^{uT_k}) = \mathbf{E}(e^{u\sum_{n=1}^{k}\tau_n}) = \mathbf{E}(\prod_{n=1}^{k} e^{u\tau_n}) = (2e^{-2u}-1)^{-k} \qquad (u < \log\sqrt{2},\ k \geqq 0). \tag{27}$$

We also need the moment generating function of the centered and "normalized" random variable $(T_k - 4k)/\sqrt{8}$, whose expectation is 0 and variance is k:

$$\mathbf{E}(e^{u(T_k-4k)/\sqrt{8}}) = e^{-4ku/\sqrt{8}}\,\mathbf{E}(e^{T_k u/\sqrt{8}}) = \left(2e^{u/\sqrt{2}} - e^{u\sqrt{2}}\right)^{-k}, \tag{28}$$

for $u < \sqrt{2}\log 2$ and $k \geqq 0$. Since (28) is less than 2^k for $u = \pm 1/2$, the exponential Chebyshev's inequality (10) implies that

$$\mathbf{P}\left\{|T_k - 4k|/\sqrt{8} \geqq t\right\} \leqq 2 \cdot 2^k e^{-t/2} \qquad (t > 0,\ k \geqq 0). \tag{29}$$

4. From random walks to the Wiener process: "twist and shrink"

Our construction of the Wiener process is based on P. Révész's one, [6, Section 6.2], which in turn is a simpler version of F.B. Knight's [4, Section 1.3]. The advantage of this method over the several known ones is that it is very natural and elementary.

We will define a sequence of approximations to the Wiener process, each of which is a "twisted and shrunk" random walk, a refinement of the previous one. It will be shown that this sequence converges to a process having the properties characterizing the Wiener process.

Imagine that we observe a particle undergoing Brownian motion. In the first experiment we observe the particle exclusively when it hits points with integer coordinates $j \in \mathbf{Z}$. Suppose that it happens exactly at the consecutive time instants $1, 2, \ldots$. To model the graph of the particle between the vertices so obtained the simplest idea is to join them by straight line segments like in Figure 1. Therefore the first approximation is

$$B_0(t) = S_0(t) = S(t),$$

where $t \geqq 0$ real and $S(t)$ is a random walk defined by (1) and (2).

Suppose that in the second experiment we observe the particle when it hits points with coordinates $j/2$ ($j \in \mathbf{Z}$), in the third experiment when it hits points with coordinates $j/2^2$ ($j \in \mathbf{Z}$), etc. To model the second experiment one idea is to take a second random walk $S_1(t)$, independent of the first one, and shrink it.

Then the first problem that arises is the relationship between the time and space scales: if one wants to compress the length of a step into half, how much one has to compress the time needed for one step to preserve the essential properties of a random walk. Here we recall that by (3), the square root of the average squared distance of the random walk from the origin after time n is $\sqrt{n}$. So shrinking the random walk so that there are n steps in one time unit, each step should have a length $1/\sqrt{n}$. This way after one time unit the square root of the average squared distance of the walk from the origin will be around one spatial unit, like in the case of the original random walk. It means that compressing the length of one step into $1/2$ (or in general: $1/2^m, m = 1, 2, \ldots$) one has to compress the time needed for one step into $1/2^2$ (in general: $1/2^{2m}$).

The second problem is that sample-functions of $B_0(t)$ and of a shrunk version of an independent $S_1(t)$ have nothing to do with each other, the second is not being a refinement of the first in general. For example, if $B_0(1) = 1$, then it is equally likely that the first integer the shrunk version of $S_1(t)$ hits is $+1$ or -1.

Hence before shrinking we want to modify $S_1(t)$ so that it hits even integers $2j$ ($j \in \mathbf{Z}$) (counting the next one only if it is different from the previous one) in exactly the same order as $S_0(t)$ hits the corresponding integers $j \in \mathbf{Z}$. For example, if $S_0(1) = 1$ and $S_0(2) = 2$, then the first even integer $S_1(t)$ hits should be 2 and the next one (different from 2) should be 4. Thus if $S_1(t)$ hits the first even integer at time $T_1(1)$ and $S_1(T_1(1))$ happens to be -2, we will reflect every step $X_1(k)$ of $S_1(t)$ for $0 < k \leqq T_1(1)$. This way we get a modified random walk $\tilde{S}_1(t)$ up to time $T_1(1)$ so that $\tilde{S}_1(T_1(1)) = 2$. Then we continue similarly up to time $T_1(2)$: if the (already modified) walk hit 0 at time $T_1(2)$ (instead of 4), then we would reflect the steps $X_1(k)$ for $T_1(1) < < k \leqq T_1(2)$. This modification process, which we will call "twisting", ensures that the next approximation will always be a refinement of the previous one.

Now let us see the construction in detail. It begins with a sequence of independent random walks $S_0(t), S_1(t), \ldots$. That is, for each $m \geqq 0$,

$$S_m(0) = 0, \quad S_m(n) = X_m(1) + X_m(2) + \cdots + X_m(n) \quad (n \geqq 1), \tag{30}$$

where $X_m(k)$ $(m \geqq 0, k \geqq 1)$ is a double array of independent, identically distributed random variables such that

$$\mathbf{P}\{X_m(k)=1\} = \mathbf{P}\{X_m(k)=-1\} = 1/2. \tag{31}$$

First we possibly modify $S_1(t), S_2(t), \ldots$ one-by-one, using the "twist" method to obtain a sequence of *not* independent random walks $\tilde{S}_1(t), \tilde{S}_2(t), \ldots$, each of which is a refinement of the former one. Second, by shrinking we get a sequence $B_1(t), B_2(t), \ldots$ approximating the Wiener process.

In accordance with the notation in (19), for $m \geqq 1$, S_m hits even integers (different from the previous one) exclusively at the random time instants

$$T_m(0) = 0, \quad T_m(k) = \tau_m(1) + \tau_m(2) + \cdots + \tau_m(k) \quad (k \geqq 1).$$

Each random variable $T_m(k)$ has the same distribution as $T(k) = T_k$ above, see (20) and (21). That is, $T_m(k)$ is the double of a negative binomial random variable, with expectation $4k$ and variance $8k$.

Now we define a suitable sequence of "twisted" random walks $\tilde{S}_m(t)$ $(m \geqq 1)$ recursively, using $\tilde{S}_{m-1}(t)$, starting with

$$\tilde{S}_0(t) = S_0(t) \qquad (t \geqq 0).$$

First we set

$$\tilde{S}_m(0) = 0.$$

Then for $k = 0, 1, \ldots$ successively and for every n such that $T_m(k) < n \leqq T_m(k+1)$, we take (Figures 2-4)

$$\tilde{X}_m(n) = \begin{cases} X_m(n) & \text{if } \; S_m(T_m(k+1)) - S_m(T_m(k)) = 2\tilde{X}_{m-1}(k+1); \\ -X_m(n) & \text{otherwise.} \end{cases} \tag{32}$$

and

$$\tilde{S}_m(n) = \tilde{S}_m(n-1) + \tilde{X}_m(n). \tag{33}$$

Observe that the stopping times $\tilde{T}_m(k)$ corresponding to $\tilde{S}_m(t)$ are the same as the original ones $T_m(k)$ $(m \geqq 0, k \geqq 0)$.

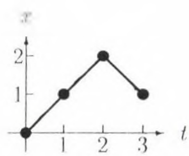

Fig. 2. $B_0(t;\omega) = S_0(t;\omega)$.

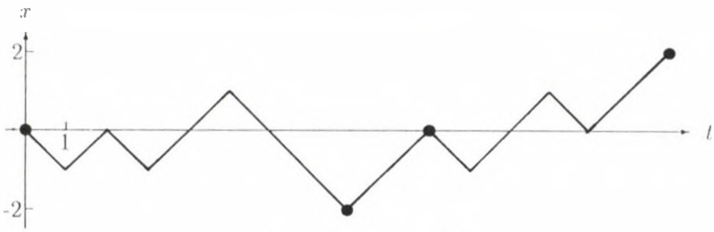

Fig. 3. $S_1(t;\omega)$.

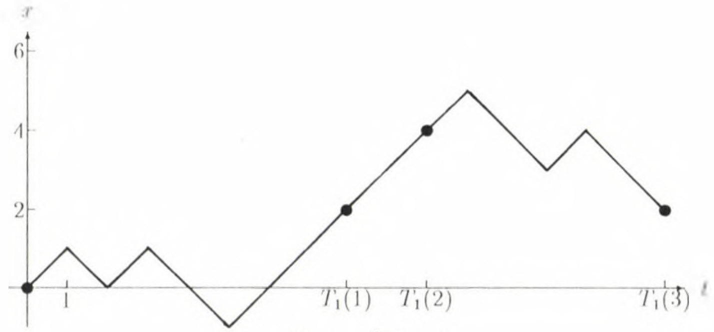

Fig. 4. $\tilde{S}_1(t;\omega)$.

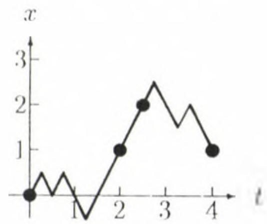

Fig. 5. $B_1(t;\omega)$.

LEMMA 1. *For each $m \geqq 0$, $\tilde{S}_m(t)$ $(t \geqq 0)$ is a random walk, that is, $\tilde{X}_m(1), \tilde{X}_m(2), \ldots$ is a sequence of independent, identically distributed random variables such that*

$$\mathbf{P}\left\{\tilde{X}_m(n) = 1\right\} = \mathbf{P}\left\{\tilde{X}_m(n) = -1\right\} = 1/2 \qquad (n \geqq 1). \tag{34}$$

PROOF. We proceed by induction over $m \geqq 0$. For $m = 0$, $\tilde{S}_0(t) = S_0(t)$, a random walk by definition. So assume that $\tilde{S}_{m-1}(t)$ is a random walk, where $m \geqq 1$, and see if it implies that $\tilde{S}_m(t)$ is a random walk, too.

It is enough to show that for any $n \geqq 1$ and any $\epsilon_j = \pm 1$ $(j = 1, \ldots, n)$ we have

$$\mathbf{P}\left\{\tilde{X}_m(1) = \epsilon_1, \ldots, \tilde{X}_m(n-1) = \epsilon_{n-1}, \tilde{X}_m(n) = \epsilon_n\right\} = 1/2^n. \tag{35}$$

Set the events $A_{m,r} = \left\{\tilde{X}_m(j) = \epsilon_j,\ 1 \leqq j \leqq r\right\}$ for $1 \leqq r \leqq n$ ($A_{m,0}$ is the sure event by definition) and the random variables $\Delta S^*_{m,k} = S_m(T_m(k+1)) - S_m(T_m(k))$ for $k \geqq 0$. The event $A_{m,n-1}$ determines the greatest integer $k \geqq 0$ such that $T_m(k) \leqq n-1$; let us denote this value by κ. By (32),

$$\mathbf{P}\{A_{m,n}\} = \sum_{\alpha=\pm 1} \mathbf{P}\left\{A_{m,n-1};\, X_m(n) = \alpha\epsilon_n;\ \Delta S^*_{m,\kappa} = \alpha 2\tilde{X}_{m-1}(\kappa+1)\right\}. \tag{36}$$

The event $A_{m,n-1}$ can be written here as $B_{m,n-1}C_{m,n-1}$, where

$$B_{m,n-1} = \left\{\tilde{X}_m(j) = \epsilon_j,\ 1 \leqq j \leqq T_m(\kappa)\right\},$$
$$C_{m,n-1} = \{X_m(j) = \alpha\epsilon_j,\ T_m(\kappa)+1 \leqq j \leqq n-1\}.$$

Definition (32) shows that $B_{m,n-1}$ is determined by $\tilde{X}_{m-1}(j)$ $(1 \leqq j \leqq \kappa)$ and $X_m(j)$ $(1 \leqq j \leqq T_m(\kappa))$ the values of which do not influence anything else in (36).

Then we distinguish two cases according to the parity of n.

Case 1. n is odd. Then $n-1$ is even and $S_m(T_m(\kappa)) = S_m(n-1)$. Further, let $\tau_{m,r} = \min\{j : j > 0, |S_m(r+j) - S_m(r)| = 2\}$ and $\Delta S_m(r) = S_m(r + \tau_{m,r}) - S_m(r)$ for $r \geqq 0$. Then $S_m(T_m(\kappa+1)) = S_m(n-1+\tau_{m,n-1})$ and $\Delta S^*_{m,\kappa} = \Delta S_m(n-1)$. These and the argument above shows that in (36) $A_{m,n-1}$ is independent of the other terms. Consequently, (36) simplifies as

$$\mathbf{P}\{A_{m,n}\} = 2\mathbf{P}\{A_{m,n-1}\}\frac{1}{2}\sum_{\beta=\pm 1}\mathbf{P}\{X_m(n) = \epsilon_n;\ \Delta S_m(n-1) = 2\beta\}, \tag{37}$$

since the value of α is immaterial and $\mathbf{P}\left\{\tilde{X}_{m-1}(\kappa+1) = \beta\right\} = 1/2$, independently of everything else here.

Finally, (17) and (18) can be applied to (37):

$$\begin{aligned}\mathbf{P}\{A_{m,n}\} &= \mathbf{P}\{A_{m,n-1}\} \times \\ &\times \sum_{\beta=\pm 1}\mathbf{P}\{\Delta S_m(n-1) = 2\beta \mid X_m(n) = \epsilon_n\}\ \mathbf{P}\{X_m(n) = \epsilon_n\} \\ &= \mathbf{P}\{A_{m,n-1}\}\left(\frac{3}{4}+\frac{1}{4}\right)\frac{1}{2} = \frac{1}{2}\mathbf{P}\{A_{m,n-1}\},\end{aligned}$$

independently of ϵ_n.

Case 2. n is even. Then $n-2$ is even and the argument in Case 1 could be repeated with $n-2$ in place of $n-1$, with the only exception that in (36) we have an additional term $\bar{X}_m(n-1) = \alpha X_m(n-1)$. Then instead of (37) we arrive at

$$\begin{aligned}
&\mathbf{P}\{A_{m,n}\} \\
&= \mathbf{P}\{A_{m,n-2}\} \sum_{\beta=\pm 1} \mathbf{P}\{X_m(n-1) = \epsilon_{n-1}, X_m(n) = \epsilon_n;\ \Delta S_m(n-2) = 2\beta\} \\
&= \mathbf{P}\{A_{m,n-2}\} \frac{1}{2^2} \sum_{\beta=\pm 1} \mathbf{P}\{\Delta S_m(n-2) = 2\beta \mid X_m(n-1) = \epsilon_{n-1}, X_m(n) = \epsilon_n\}.
\end{aligned} \tag{38}$$

The conditional probability in (38) is

$$\begin{array}{llll}
1 & \text{if } \beta = \epsilon_{n-1} = \epsilon_n; & 0 & \text{if } -\beta = \epsilon_{n-1} = \epsilon_n; \\
1/2 & \text{if } \beta = \epsilon_{n-1} = -\epsilon_n; & 1/2 & \text{if } -\beta = \epsilon_{n-1} = -\epsilon_n.
\end{array}$$

Thus the sum in (38) becomes

$$1 + 0 = 1 \text{ if } \epsilon_{n-1} = \epsilon_n; \ (1/2) + (1/2) = 1 \text{ if } \epsilon_{n-1} = -\epsilon_n.$$

In other words, the value of the sum in (38) is 1, independently of ϵ_{n-1} and ϵ_n.

In summary, $\mathbf{P}\{A_{m,n}\} = \frac{1}{2}\mathbf{P}\{A_{m,n-1}\}$ if n is odd and $\mathbf{P}\{A_{m,n}\} = \frac{1}{4}\mathbf{P}\{A_{m,n-2}\}$ if n is even. Since $\mathbf{P}\{A_{m,0}\} = 1$, (35) follows. □

We mention that an other possibility to prove Lemma 1 is to introduce the random variables $Z_k = \frac{1}{2}\Delta S^*_{m,k-1}\tilde{X}_{m-1}(k)$ for $k \geqq 1$. It can be shown that $Z_1, Z_2, \ldots$ is a sequence of independent and identically distributed random variables, $\mathbf{P}\{Z_k = 1\} = \mathbf{P}\{Z_k = -1\} = 1/2$, and this sequence is independent of the sequence $X_m(1), X_m(2), \ldots$ as well. Then we have $\bar{X}_m(n) = Z_k X_m(n)$ for each n such that $T_m(k-1) < n \leqq T_m(k)$ $(k \geqq 1)$ and this implies (35).

The main property that was aimed when we introduced the "twist" method easily follows from (32) and (35):

$$\begin{aligned}
\bar{S}_m(T_m(k)) &= \sum_{j=1}^{k} \bar{S}_m(T_m(j)) - \bar{S}_m(T_m(j-1)) \\
&= \sum_{j=1}^{k} 2\tilde{X}_{m-1}(j) = 2\tilde{S}_{m-1}(k),
\end{aligned} \tag{39}$$

for any $m \geqq 1$ and $k \geqq 0$.

Now the second step of the approximation comes: "shrinking". As was discussed above, at the mth approximation the length of one step should be

$1/2^m$ and the time needed for a step should be $1/2^{2m}$ (Figure 5). So we define the *mth approximation of the Wiener process* by

$$B_m\left(\frac{t}{2^{2m}}\right) = \frac{1}{2^m}\tilde{S}_m(t) \qquad (t \geqq 0, m \geqq 0), \tag{40}$$

or equivalently, $B_m(t) = 2^{-m}\tilde{S}_m(t2^{2m})$. Basically, $B_m(t)$ is a model of Brownian motion on the set of points $x = j/2^m$ $(j \in \mathbf{Z})$.

Now (39) becomes the following *refinement property*:

$$B_m\left(\frac{T_m(k)}{2^{2m}}\right) = B_{m-1}\left(\frac{k}{2^{2(m-1)}}\right), \tag{41}$$

for any $m \geqq 1$ and $k \geqq 0$.

The remaining part of this section is devoted to showing the convergence of the sequence $B_m(t)$ $(m = 0, 1, 2, \ldots)$, and that the limiting process has the characterizing properties of the Wiener process. In proving these our basic tools will be some relatively simple, but powerful observations.

First, often in the sequel the following crude, but still efficient estimate will be applied:

$$\mathbf{P}\left\{\max_{1 \leqq j \leqq N} Z_j \geqq t\right\} = \mathbf{P}\left\{\bigcup_{j=1}^{N} \{Z_j \geqq t\}\right\} \leqq \sum_{j=1}^{N} \mathbf{P}\{Z_j \geqq t\}, \tag{42}$$

which is valid for arbitrary random variables Z_j and real number t.

The proofs of Lemmas 3 and 4 below essentially consist of the application of the following large deviation type estimate fulfilled by S_n and $(T_k - 4k)/\sqrt{8}$ according to Theorems 1(b) and 2(b). The previously mentioned exponential Chebyshev's inequalities (12) and (29) will also be used. Note that in the next lemma we have $a = 2$ and $b = 1$ for S_n in (12) and $a = 2$ and $b = 1/2$ for $(T_k - 4k)/\sqrt{8}$ in (19).

Lemma 2. *Suppose that for $j \geqq 0$, we have $\mathbf{E}(Z_j) = 0$, $\mathbf{Var}(Z_j) = j$, and with some $a > 0$ and $b > 0$,*

$$\mathbf{P}\{|Z_j| \geqq t\} \leqq 2a^j e^{-bt} \qquad (t > 0)$$

(exponential Chebyshev-type inequality).

Assume as well that there exists a $j_0 > 0$ such that for any $j \geqq j_0$,

$$\mathbf{P}\left\{|Z_j|/\sqrt{j} \geqq x_j\right\} \leqq e^{-x_j^2/2},$$

whenever $x_j \to \infty$ and $x_j = o(j^{1/6})$ as $j \to \infty$ (large deviation type inequality).

Then for any $C > 1$,

$$\mathbf{P}\left\{\max_{0 \leqq j \leqq N} |Z_j| \geqq \sqrt{2CN \log N}\right\} \leqq \frac{2}{N^{C-1}}, \tag{43}$$

if N is large enough, $N \geqq N_0(C)$.

PROOF. The maximum in (43) can be handled by the crude estimate (42). Divide the resulting sum into two parts: one that can be estimated by a large deviation type inequality, and an other that will be estimated using exponential Chebyshev's inequality. For the large deviation part x_j will be $\sqrt{2C \log N}$. Since $j \leqq N$, $j \to \infty$ implies that $N \to \infty$, and then $x_j \to \infty$ as well. If $j \geqq \log^4 N$, then the condition $x_j = o(j^{1/6})$ holds, too, and the large deviation type inequality is applicable. Thus

$$\mathbf{P}\left\{\max_{0 \leqq j \leqq N} |Z_j| \geqq \sqrt{2CN \log N}\right\}$$
$$\leqq \sum_{j=0}^{\lfloor \log^4 N \rfloor} 2a^j \exp\left(-b\sqrt{2CN \log N}\right) + \sum_{j=\lfloor \log^4 N \rfloor}^{N} \mathbf{P}\left\{|Z_j|/\sqrt{j} \geqq \sqrt{2C \log N}\right\}$$
$$\leqq \frac{2a}{a-1} \exp\left(\log a \ \log^4 N - b\sqrt{2CN \log N}\right) + N \exp(-C \log N) \leqq 2 \ N^{1-C}$$

if $C > 1$ and $N \geqq N_0(C)$. ($\lfloor x \rfloor$ denotes the greatest integer less than or equal to x.) □

Note that the lemma and its proof are valid even when N is not an integer. Here and afterwards we use the convention that if the upper limit of a sum is a real value N, then the sum goes until $\lfloor N \rfloor$.

We mention that both inequalities among the assumptions of the previous Lemma 2 hold for partial sums Z_j of any sequence of independent and identically distributed random variables with expectation 0, variance 1 and a moment generating function which is finite for some $\pm u_0$. The fact that an exponential Chebyshev-type inequality should hold then can be seen from (10) and (11), while the large deviation type estimate is shown to hold e.g. in [2, Section XVI,6].

The *first Borel–Cantelli lemma* [2, Section VIII,3] is also an important tool, stating that if there is given an infinite sequence $A_1, A_2, \ldots$ of events such that $\sum_{m=1}^{\infty} \mathbf{P}\{A_m\}$ is finite, then *with probability* 1 only finitely many of the events occur. Or with an other widely used terminology: *almost surely* only finitely many of them will occur.

Now turning to the convergence proof, as the first step, it will be shown that the time instants $T_{m+1}(k)/2^{2(m+1)}$ will get arbitrarily close to the time instants $k/2^{2m} = 4k/2^{2(m+1)}$ as $m \to \infty$. By (41), this means that the next approximation not only visits the points $x = j/2^m$ ($j \in \mathbf{Z}$) in the same order, but the corresponding time instants will get arbitrarily close to each other as $m \to \infty$. Remember that by (20) and (21), $T_m(k)$ is the double of a negative binomial random variable, with expectation $4k$ and variance $8k$. Here Lemma 2 will be applied to $(T_m(k) - 4k)/\sqrt{8}$ with $N = K2^{2m}$. So $\log N = \log K + (2\log 2)m \leqq 1.5m$ if m is large enough, $m \geqq m_0(K)$, and then $\sqrt{2CN\log N} \leqq \sqrt{3CKm}2^m$.

LEMMA 3. (a) *For any $C > 1$, $K > 0$, and for any $m \geqq m_0(C,K)$ we have*

$$\mathbf{P}\left\{\max_{0\leqq k/2^{2m}\leqq K} |T_{m+1}(k) - 4k| \geqq \sqrt{24CKm}2^m\right\} < 2(K2^{2m})^{1-C}. \tag{44}$$

(b) *For any $K > 0$,*

$$\max_{0\leqq k/2^{2m}\leqq K}\left|\frac{T_{m+1}(k)}{2^{2(m+1)}} - \frac{k}{2^{2m}}\right| < \sqrt{2Km}2^{-m} \tag{45}$$

with probability 1 *for all but finitely many m.*

PROOF. (a) (44) is a direct consequence of Lemma 2.

(b) Take for example $C = 4/3$ in (a) and define the following events for $m \geqq 0$:

$$A_m = \left\{\max_{0\leqq k/2^{2m}\leqq K} |T_{m+1}(k) - 4k| \geqq \sqrt{32Km}2^m\right\}.$$

By (44), for $m \geqq m_0(C,K)$, $\mathbf{P}\{A_m\} < 2(K2^{2m})^{-1/3}$. Then $\sum_{m=0}^{\infty}\mathbf{P}\{A_m\} < < \infty$. Hence the Borel–Cantelli lemma implies that with probability 1, only finitely many of the events A_m occur. That is, almost surely for all but finitely many m we have

$$\max_{0\leqq k/2^{2m}\leqq K} |T_{m+1}(k) - 4k| < \sqrt{32Km}2^m.$$

This inequality is equivalent to (45). □

It seems to be important to emphasize a "weakness" of a statement like the one in Lemma 3(b): we use the phrase "all but finitely many m" to indicate that the statement holds for every $m \geqq m_0(\omega)$, where $m_0(\omega)$ may depend on the specific point ω of the sample space. In other words, one has no common, uniform lower bound for m in general.

Next we want to show that for any $j \geqq 1$, $B_{n+j}(t)$ will be arbitrarily close to $B_n(t)$ as $n \to \infty$. Here again Lemma 2 will be applied, this time to a random walk S_r, with a properly chosen N' and C' (instead of $N = K2^{2m}$ and C). Although the proof will be somewhat long, its basic idea is simple. Since $B_{m+1}(T_{m+1}(k)/2^{2(m+1)}) = B_m(k/2^{2m})$ by (39), and the difference of the corresponding time instants here approaches zero fast as $m \to \infty$ by (45), one can show that $B_m(t)$ and its refinement $B_{m+1}(t)$ will get very close to each other, too.

The following elementary fact that we need in the proof is discussed before stating the lemma:

$$\sum_{m=n}^{\infty} m 2^{-m/2} = (1/\sqrt{2}) \sum_{m=n}^{\infty} m \left(1/\sqrt{2}\right)^{m-1} < 4n2^{-n/2}, \tag{46}$$

for $n \geqq 15$. This can be shown by a routine application of power series:

$$\sum_{m=n}^{\infty} m x^{m-1} = \frac{d}{dx} \sum_{m=n}^{\infty} x^m = \frac{d}{dx}\left(\frac{x^n}{1-x}\right) = n x^{n-1}\left(\frac{1}{1-x} + \frac{x}{n(1-x)^2}\right).$$

Substituting $x = 1/\sqrt{2}$, one gets (46) for $n \geqq 15$.

LEMMA 4. (a) *For any* $C \geqq 3/2$, $K > 0$, *and for any* $n \geqq n_0(C, K)$ *we have*

$$\begin{aligned} \mathbf{P}&\left\{\max_{0 \leqq k/2^{2n} \leqq K} |B_{n+1}(T_{n+1}(k)/2^{2(n+1)}) - B_{n+1}(k/2^{2n})| \geqq (1/8)n2^{-n/2}\right\} \\ &\leqq 3(K2^{2n})^{1-C} \end{aligned} \tag{47}$$

and

$$\mathbf{P}\left\{\max_{0 \leqq t \leqq K} |B_{n+j}(t) - B_n(t)| \geqq n2^{-n/2} \text{ for some } j \geqq 1\right\} < 6(K2^{2n})^{1-C}. \tag{48}$$

(b) *For any* $K > 0$,

$$\max_{0 \leqq t \leqq K} |B_{n+j}(t) - B_n(t)| < n2^{-n/2}, \tag{49}$$

with probability 1 *for all* $j \geqq 1$ *and for all but finitely many* n.

PROOF. Let us consider first the difference between two consecutive approximations, $\max_{0 \leqq t \leqq K} |B_{m+1}(t) - B_m(t)|$. The maximum over real values t can be approximated by the maximum over dyadic rational numbers $k/2^{2m}$. This is true because any sample-function $B_m(t;\omega)$ is a broken line such that, by (40), the magnitude of the increment between two consecutive points $k/2^{2m}$ and $(k+1)/2^{2m}$ is equal to 2^{-m}. Thus, taking the integer $t_m = \lfloor t2^{2m} \rfloor$ for each $t \in [0, K]$, one has $t_m/2^{2m} \leqq t < (t_m+1)/2^{2m}$ and so $4t_m/2^{2(m+1)} \leqq t < (4t_m+4)/2^{2(m+1)}$. So we get $|B_m(t) - B_m(t_m/2^{2m})| < 2^{-m}$ and $\left|B_{m+1}(t) - B_{m+1}\left(4t_m/2^{2(m+1)}\right)\right| \leqq 4 \cdot 2^{-(m+1)} = 2 \cdot 2^{-m}$. Hence

$$\begin{aligned} &\max_{0 \leqq t \leqq K} |B_{m+1}(t) - B_m(t)| \\ &\leqq \max_{0 \leqq k/2^{2m} \leqq K} \left|B_{m+1}\left(4k/2^{2(m+1)}\right) - B_m\left(k/2^{2m}\right)\right| + 3 \cdot 2^{-m}. \end{aligned}$$

Moreover, by (41) and (40) we have

$$
\begin{aligned}
&B_{m+1}\left(4k/2^{2(m+1)}\right) - B_m\left(k/2^{2m}\right) \\
(50)\qquad &= B_{m+1}\left(4k/2^{2(m+1)}\right) - B_{m+1}\left(T_{m+1}(k)/2^{2(m+1)}\right) \\
&= 2^{-(m+1)}\tilde{S}_{m+1}(4k) - 2^{-(m+1)}\tilde{S}_{m+1}(T_{m+1}(k)).
\end{aligned}
$$

Thus

$$
\begin{aligned}
&\mathbf{P}\left\{\max_{0\leqq t\leqq K}|B_{m+1}(t)-B_m(t)| \geqq (1/4)m2^{-m/2}\right\} \\
&\leqq \mathbf{P}\left\{\max_{0\leqq k/2^{2m}\leqq K}\left|B_{m+1}\left(4k/2^{2(m+1)}\right) - B_m\left(k/2^{2m}\right)\right| \geqq (1/8)m2^{-m/2}\right\} \\
(51)\qquad &= \mathbf{P}\left\{\max_{0\leqq k\leqq K2^{2m}}|\tilde{S}_{m+1}(4k)-\tilde{S}_{m+1}(T_{m+1}(k))| \geqq (1/4)m2^{m/2}\right\}
\end{aligned}
$$

if m is large enough.

By Lemma 3, the probability of the event

$$
A_m = \left\{\max_{0\leqq k\leqq K2^{2m}}|T_{m+1}(k)-4k| \geqq \sqrt{24CKm}2^m\right\}
$$

is very small for m large. Therefore we divide the last expression in (51) into two parts according to A_m and A_m^c (the complement of A_m):

$$
\begin{aligned}
&\mathbf{P}\left\{\max_{0\leqq t\leqq K}|B_{m+1}(t)-B_m(t)| \geqq (1/4)m2^{-m/2}\right\} \\
&\leqq \mathbf{P}\left\{A_m^c;\max_{0\leqq k\leqq K2^{2m}}\left|\tilde{S}_{m+1}(4k)-\tilde{S}_{m+1}(T_{m+1}(k))\right| \geqq (1/4)m2^{m/2}\right\} + \mathbf{P}\{A_m\} \\
(52)\qquad &\leqq \sum_{k=1}^{K2^{2m}}\mathbf{P}\left\{\max_{\{j:\,|j-4k|<\sqrt{24CKm}2^m\}}\left|\tilde{S}_{m+1}(4k)-\tilde{S}_{m+1}(j)\right| \geqq (1/4)m2^{m/2}\right\} \\
&\qquad +2(K2^{2m})^{1-C},
\end{aligned}
$$

where the crude estimate (42) and Lemma 3(a) were used.

Now apply Lemma 2 to $\tilde{S}_{m+1}(j)-\tilde{S}_{m+1}(4k)$ here, with suitably chosen N' and C'. For k fixed and $j>4k$, $\tilde{S}_{m+1}(j)-\tilde{S}_{m+1}(4k)=\sum_{r=4k+1}^{j}X_{m+1}(r)$ is a random walk of the form $S(j-4k)$. (The case $j<4k$ is symmetric.) Since $|j-4k|<\sqrt{24CKm}2^m$, N' is taken as $\sqrt{24CKm}2^m$. (So N' is roughly

$\sqrt{N}$, where $N = K2^{2m}$.) Then $\log N' = (1/2)\log(24CKm) + (\log 2)m \leqq m$ if m is large enough, depending on C and K. So

$$\sqrt{2C'N'\log N'} \leqq \sqrt{2C'm\sqrt{24CKm}2^m} \leqq (1/4)m2^{m/2},$$

if m is large enough, depending on C, C', and K. Then it follows by Lemma 2 that

$$\mathbf{P}\left\{\max_{0\leqq r\leqq\sqrt{24CKm}2^m} |S(r)| \geqq (1/4)m2^{m/2}\right\} \leqq 2(\sqrt{24CKm}2^m)^{1-C'}. \tag{53}$$

The second term of the error probability in (52) is $2(K2^{2m})^{1-C} = 2N^{1-C}$, while (53) indicates that the first term is at most $K2^{2m}2\cdot 2(\sqrt{24CKm}2^m)^{1-C'}$ $\leqq N(\sqrt{N})^{1-C'}$ if $C' > 1$ and m is large enough. To make the two error terms to be of the same order, choose $1 + (1-C')/2 = 1 - C$, i.e. $C' = 2C+1$. Thus (52) becomes

$$\mathbf{P}\left\{\max_{0\leqq t\leqq K} |B_{m+1}(t) - B_m(t)| \geqq (1/4)m2^{-m/2}\right\} \leqq 3(K2^{2m})^{1-C},$$

for any m large enough, depending on C and K. Comparing this to (50) and (51) one obtains (47).

By (46), $\max_{0\leqq t\leqq K} |B_{m+1}(t) - B_m(t)| < (1/4)m2^{-m/2}$ for all $m \geqq n \geqq 15$ would imply that

$$\max_{0\leqq t\leqq K} |B_{n+j}(t) - B_n(t)| = \max_{0\leqq t\leqq K}\left|\sum_{m=n}^{n+j-1} B_{m+1}(t) - B_m(t)\right|$$
$$\leqq \sum_{m=n}^{n+j-1} \max_{0\leqq t\leqq K} |B_{m+1}(t) - B_m(t)| < \sum_{m=n}^{\infty} (1/4)m2^{-m/2} < n2^{-n/2},$$

for any $j \geqq 1$. So we conclude that

$$\mathbf{P}\left\{\max_{0\leqq t\leqq K} |B_{n+j}(t) - B_n(t)| \geqq n2^{-n/2} \text{ for some } j \geqq 1\right\}$$
$$\leqq \sum_{m=n}^{\infty} \mathbf{P}\left\{\max_{0\leqq t\leqq K} |B_{m+1}(t) - B_m(t)| \geqq (1/4)m2^{-m/2}\right\}$$
$$\leqq \sum_{m=n}^{\infty} 3(K2^{2m})^{1-C} = 3(K2^{2n})^{1-C}\frac{1}{1-2^{2(1-C)}} < 6(K2^{2n})^{1-C}$$

if $C \geqq 3/2$ (say), for any $n \geqq n_0(C,K)$. This proves (48).

The statement in (b) follows from (48) by the Borel–Cantelli lemma, as in the proof of Lemma 3. □

Now we are ready to establish the existence of the Wiener process, which is a continuous model of Brownian motion. An important consequence of (49) is that the difference between the Wiener process and $B_n(t)$ is smaller than a constant multiple of $\log N/\sqrt[4]{N}$, where $N = K2^{2n}$, see (55) below.

THEOREM 3. *As $n \to \infty$, with probability 1 (that is, for almost all $\omega \in \Omega$) and for all $t \in [0,\infty)$ the sample-functions $B_n(t;\omega)$ converge to a sample-function $W(t;\omega)$ such that*

(i) *$W(0;\omega) = 0$, $W(t;\omega)$ is a continuous function of t on the interval $[0,\infty)$;*

(ii) *for any $0 \leqq s < t$, $W(t) - W(s)$ is a normally distributed random variable with expectation 0 and variance $t - s$;*

(iii) *for any $0 \leqq s < t \leqq u < v$, the increments $W(t) - W(s)$ and $W(v) - W(u)$ are independent random variables.*

By definition, $W(t)$ $(t \geqq 0)$ is called the **Wiener process**.

Further, we have the following estimates for the difference of the Wiener process and its approximations.

(a) *For any $C \geqq 3/2$, $K > 0$, and for any $n \geqq n_0(C,K)$ we have*

$$\mathbf{P}\left\{\max_{0\leqq t\leqq K} |W(t) - B_n(t)| \geqq n2^{-n/2}\right\} \leqq 6(K2^{2n})^{1-C}. \tag{54}$$

(b) *For any $K > 0$,*

$$\max_{0\leqq t\leqq K} |W(t) - B_n(t)| \leqq n2^{-n/2}, \tag{55}$$

with probability 1 for all but finitely many n.

PROOF. Lemma 4(b) shows that for almost all $\omega \in \Omega$, the sequence $B_n(t;\omega)$ converges for any $t \geqq 0$ as $n \to \infty$. Let us denote the limit by $W(t;\omega)$. On a probability zero ω-set the limit possibly does not exist, there one can define $W(t;\omega) = 0$ for any $t \geqq 0$. Since $B_n(0;\omega) = 0$ for any n, it follows that $W(0;\omega) = 0$ for any $\omega \in \Omega$.

Taking $j \to \infty$ in (48), (54) follows. By (49), the convergence of $B_n(t)$ is uniform on any bounded interval $[0,K]$, more exactly, for any $K > 0$ we have (55) with probability 1. Textbooks on advanced calculus, like W. Rudin's [7, Section 7.12] show that the limit function of a uniformly convergent sequence of continuous functions is also continuous. This proves (i).

Now we turn to the proof of (ii). Take arbitrary $t > s \geqq 0$ and x real. With $K > t$ fixed, (54) shows that for any $\delta > 0$ there exists an $n \geqq n_0(C,K)$ such that

$$\mathbf{P}\left\{\max_{0\leqq u\leqq K} |W(u) - B_n(u)| \geqq \delta\right\} < \delta. \tag{56}$$

Since

$$\mathbf{P}\{W(t)-W(s)\leqq x\}=$$
$$\mathbf{P}\{B_n(t)-B_n(s)\leqq x-(W(t)-B_n(t))+(W(s)-B_n(s))\},$$

(56) implies that

$$\begin{aligned}\mathbf{P}\{B_n(t)-B_n(s)\leqq x-2\delta\}-2\delta\leqq \mathbf{P}\{W(t)-W(s)\leqq x\}\\ \leqq \mathbf{P}\{B_n(t)-B_n(s)\leqq x+2\delta\}+2\delta.\end{aligned} \tag{57}$$

This indicates that the distribution function of $W(t)-W(s)$ can be eventually obtained from the distribution function of

$$B_n(t)-B_n(s)=2^{-n}\bar{S}_n(2^{2n}t)-2^{-n}\bar{S}_n(2^{2n}s). \tag{58}$$

Take the non-negative integers $j_n=\lfloor 2^{2n}t\rfloor$ and $i_n=\lfloor 2^{2n}s\rfloor$, $j_n\geqq i_n$. Then (58) differs from

$$2^{-n}(\bar{S}_n(j_n)-\bar{S}_n(i_n))=2^{-n}\sum_{k=i_n+1}^{j_n}\bar{X}_k \tag{59}$$

by an error not more than $2\cdot 2^{-n}<\delta$. (We can assume that n was chosen so.) Also, j_n-i_n differs from $2^{2n}(t-s)$ by at most 1. In particular, $j_n-i_n\to\infty$ as $n\to\infty$.

If n is large enough (we can assume again that n was chosen so), by Theorem 1(a), for any fixed real x' we have

$$\Phi(x')-\delta\leqq \mathbf{P}\left\{\frac{1}{\sqrt{j_n-i_n}}\sum_{k=i_n+1}^{j_n}\bar{X}_k\leqq x'\right\}\leqq\Phi(x')+\delta. \tag{60}$$

Here $\sqrt{j_n-i_n}$ can be approximated by $2^n\sqrt{t-s}$ if n is large enough,

$$1-\delta<\sqrt{\frac{j_n-i_n-1}{j_n-i_n}}\leqq\frac{2^n\sqrt{t-s}}{\sqrt{j_n-i_n}}\leqq\sqrt{\frac{j_n-i_n+1}{j_n-i_n}}<1+\delta. \tag{61}$$

Combining formulae (58)–(61) we obtain that

$$\begin{aligned}\Phi\left((1-\delta)\frac{x}{\sqrt{t-s}}-\delta\right)-\delta\leqq \mathbf{P}\{B_n(t)-B_n(s)\leqq x\}\\ \leqq\Phi\left((1+\delta)\frac{x}{\sqrt{t-s}}+\delta\right)+\delta.\end{aligned}$$

This shows that the distribution of $B_n(t)-B_n(s)$ is asymptotically normal with mean 0 and variance $t-s$ as $n\to\infty$. Moreover, by (57), the distribution of $W(t) - W(s)$ is *exactly* normal with mean 0 and variance $t - s$, since δ can be made arbitrarily small if n is large enough:

$$\mathbf{P}\left\{W(t) - W(s) \leqq x\right\} = \Phi\left(\frac{x}{\sqrt{t-s}}\right).$$

This proves (ii).

Finally, (iii) can be proved similarly as (ii) above. Taking arbitrary $v > u \geqq t > s \geqq 0$ and x, y real numbers,

$$\mathbf{P}\left\{W(t) - W(s) \leqq x,\ W(v) - W(u) \leqq y\right\} \tag{62}$$

can be approximated by a probability of the form

$$\mathbf{P}\left\{B_n(t) - B_n(s) \leqq x\ ,\ B_n(v) - B_n(u) \leqq y\right\}$$

arbitrarily well if n is large enough, just like in (57). In turn, like in (59), the latter can be estimated arbitrarily well by a probability of the form

$$\mathbf{P}\left\{\frac{1}{\sqrt{j_n - i_n}} \sum_{k=i_n+1}^{j_n} \tilde{X}_k \leqq x'\ ,\ \frac{1}{\sqrt{r_n - q_n}} \sum_{k=q_n+1}^{r_n} \tilde{X}_k \leqq y'\right\}, \tag{63}$$

where $i_n = \lfloor 2^{2n}s \rfloor \leqq j_n = \lfloor 2^{2n}t \rfloor \leqq q_n = \lfloor 2^{2n}u \rfloor \leqq r_n = \lfloor 2^{2n}v \rfloor$.

Since there are no common terms in the first and the second sum of (63), the two sums are independent. Thus (63) is equal to

$$\mathbf{P}\left\{\frac{1}{\sqrt{j_n - i_n}} \sum_{k=i_n+1}^{j_n} \tilde{X}_k \leqq x'\right\} \cdot \mathbf{P}\left\{\frac{1}{\sqrt{r_n - q_n}} \sum_{k=q_n+1}^{r_n} \tilde{X}_k \leqq y'\right\},$$

which can be made arbitrarily close to

$$\mathbf{P}\left\{W(t) - W(s) \leqq x\right\} \cdot \mathbf{P}\left\{W(v) - W(u) \leqq y\right\}. \tag{64}$$

Since errors in the approximations can be made arbitrarily small, (62) and (64) must agree for any real x and y. This proves (iii). $\square$

Note that properties (ii) and (iii) are often rephrased in the way that the Wiener process is a Gaussian process with independent and stationary increments. It can be proved [4, Section 1.5] that properties (i), (ii), and (iii) characterize the Wiener process. In other words, any construction to the Wiener process gives essentially the same process that was constructed above.

5. From the Wiener process to random walks

Now we are going to check whether the Wiener process as a model of Brownian motion has the properties described in the introduction of Section 4. Namely, we would want to find the sequence of shrunk random walks $B_m(k2^{-2m})$ in $W(t)$.

Let $s(1)$ be the first (random) time instant where the magnitude of the Wiener process is 1: $s(1) = \min\{s > 0 : |W(s)| = 1\}$. The continuity and increment characteristics of the Wiener process imply that $s(1)$ exists with probability 1. Clearly, each shrunk random walk $B_m(t)$ has the symmetry property that reflecting all its sample-functions to the time axis, one gets the same process. $W(t)$ as a limiting process of shrunk random walks inherits this feature. Therefore setting $X(1) = W(s(1))$, $\mathbf{P}\{X(1) = 1\} = \mathbf{P}\{X(1) = -1\} = 1/2$.

Inductively, starting with $s(0) = 0$, if $s(k-1)$ is given, define the random time instant

$$s(k) = \min\{s : s > s(k-1), |W(s) - W(s(k-1))| = 1\} \qquad (k \geqq 1).$$

As above, $s(k)$ exists with probability 1. Setting $X(k) = W(s(k)) - W(s(k-1))$, it is heuristically clear that $\mathbf{P}\{X(k) = 1\} = \mathbf{P}\{X(k) = -1\} = 1/2$, and $X(k)$ is independent of $X(1), X(2), \ldots, X(k-1)$.

This way one gets a random walk $S(k) = W(s(k)) = X(1) + X(2) + \cdots + X(k)$ $(k \geqq 0)$ from the Wiener process. Using a more technical phrase, by this method, based on *first passage times*, one can *imbed* a random walk into the Wiener process; it is a special case of the famous Skorohod imbedding, see e.g. [1, Section 13.3].

Quite similarly, one can imbed $B_m(k2^{-2m})$ into $W(t)$ for any $m \geqq 0$ by setting $s_m(0) = 0$,

(65)
$$s_m(k) = \min\left\{s : s > s_m(k-1), |W(s) - W(s_m(k-1))| = 2^{-m}\right\} \quad (k \geqq 1),$$

and $B_m(k2^{-2m}) = W(s_m(k))$ $(k \geqq 0)$.

However, instead of proving all necessary details about Skorohod imbedding briefly described above, we will define an other imbedding method better suited to our approach. It will turn out that our imbedding is essentially equivalent to the Skorohod imbedding.

Our task requires a more careful analysis of the waiting times $T_m(k)$ first. Recall the refinement property (41) of $B_m(t)$. Continuing that, we get

$$\begin{aligned} B_m(k2^{-2m}) &= B_{m+1}(2^{-2(m+1)}T_{m+1}(k)) \\ &= B_{m+2}(2^{-2(m+2)}T_{m+2}(T_{m+1}(k))) = \cdots \\ &= B_n(2^{-2n}T_n(T_{n-1}(\cdots(T_{m+1}(k))\cdots))), \end{aligned} \tag{66}$$

where $k \geqq 0$ and $n > m \geqq 0$. In other words, $B_n(t)$, $n > m$, visits the same dyadic points $k2^{-m}$ in the same order as $B_m(t)$, only the corresponding time instants can differ.

To simplify the notation, let us introduce

$$T_{m,n}(k) = T_n(T_{n-1}(\cdots(T_{m+1}(k))\cdots)) \qquad (n > m \geqq 0,\ k \geqq 0).$$

Then (66) becomes

$$B_m(k2^{-2m}) = B_n(2^{-2n}T_{m,n}(k)) \qquad (n > m \geqq 0,\ k \geqq 0). \tag{67}$$

The next lemma considers *time lags* of the form $2^{-2n}T_{m,n}(k) - k2^{-2m}$.

Note that in the proofs of the next two lemmas we make use of the following simple inequality, valid for arbitrary random variables Z_j, real numbers t_j, and events $A_j = \{Z_j \geqq t_j\}$:

$$\begin{aligned} &\mathbf{P}\{Z_j \geqq t_j \text{ for some } j \geqq 1\} = \mathbf{P}\left\{\bigcup_{j=1}^{\infty} A_j\right\} \\ &= \mathbf{P}\{A_1\} + \mathbf{P}\{A_1^c A_2\} + \cdots + \mathbf{P}\left\{A_1^c \cdots A_j^c A_{j+1}\right\} + \cdots \\ &\leqq \mathbf{P}\{Z_1 \geqq t_1\} + \sum_{j=1}^{\infty} \mathbf{P}\{Z_j < t_j, Z_{j+1} \geqq t_{j+1}\}. \end{aligned} \tag{68}$$

LEMMA 5. (a) *For any $C \geqq 3/2$, $K > 0$, and $m \geqq 0$ take the following subset of the sample space:*

(69)

$$A_m = \left\{\max_{0 \leqq k2^{-2m} \leqq K} |2^{-2n}T_{m,n}(k) - k2^{-2m}| < \sqrt{18CKm}\,2^{-m} \text{ for all } n > m\right\}.$$

Then for any $m \geqq m_0(C, K)$ we have

$$\mathbf{P}\{A_m\} \geqq 1 - 4(K2^{2m})^{1-C}. \tag{70}$$

(b) *For any $K > 0$,*

$$\max_{0 \leqq k2^{-2m} \leqq K} |2^{-2n}T_{m,n}(k) - k2^{-2m}| < \sqrt{27Km}\,2^{-m}$$

with probability 1 for all $n > m$, for all but finitely many m.

PROOF. Take any β, $1/2 < \beta < 1$, and $K' > K$; for example $\beta = 2/3$ and $K' = (4/3)K$ will do. Set $\alpha_j = 1 + \beta + \beta^2 + \cdots + \beta^j$ for $j \geqq 0$,

$$Z_n = \max_{0 \leqq k2^{-2m} \leqq K} |T_{m,n}(k) - k2^{2(n-m)}|, \ t_n = \alpha_{n-m-1}\sqrt{24CK'm}\,2^{2n-m-2},$$

and $Y_{n+1} = \max_{0 \leqq k2^{-2m} \leqq K} |T_{n+1}(T_{m,n}(k)) - 4T_{m,n}(k)|$ for $n > m \geqq 0$.

First, by Lemma 3(a),

$$\mathbf{P}\{Z_{m+1} \geqq t_{m+1}\} =$$
$$= \mathbf{P}\left\{\max_{0 \leqq k2^{-2m} \leqq K} |T_{m+1}(k) - 4k| \geqq \sqrt{24CK'm}2^m\right\} \leqq 2(K2^{2m})^{1-C}$$

if m is large enough.

Second, by the triangle inequality, $Z_{n+1} \leqq 4Z_n + Y_{n+1}$ for any $n > m$. So

$$\mathbf{P}\{Z_n < t_n, Z_{n+1} \geqq t_{n+1}\} \leqq \mathbf{P}\{Z_n < t_n, Y_{n+1} \geqq t_{n+1} - 4t_n\}.$$

If $Z_n < t_n$, then setting $j = T_{m,n}(k)$,

$$\begin{aligned} j2^{-2n} &< 2^{-2n}(k2^{2(n-m)} + t_n) = k2^{-2m} + \alpha_{n-m-1}\sqrt{24CK'm}2^{-m-2} \\ &\leqq K + 3\sqrt{2CKm}2^{-m} \leqq (4/3)K = K' \end{aligned}$$

holds for $m \geqq m_0(C, K)$, since $\alpha_r < 1/(1-\beta) = 3$ (if $\beta = 2/3$) for any $r \geqq 0$. Applying these first and Lemma 3(a) last, for $n > m \geqq m_0(C, K)$ we get that

$$\begin{aligned} &\mathbf{P}\{Z_n < t_n, Z_{n+1} \geqq t_{n+1}\} \\ &\leqq \mathbf{P}\left\{\max_{0 \leqq j2^{-2n} \leqq K'} |T_{n+1}(j) - 4j| \geqq \sqrt{24CK'm}2^n 2^{n-m}(\alpha_{n-m} - \alpha_{n-m-1})\right\} \\ &\leqq \mathbf{P}\left\{\max_{0 \leqq j \leqq K'2^{2n}} |T_{n+1}(j) - 4j| \geqq \sqrt{24CK'n}2^n\right\} \leqq 2(K'2^{2n})^{1-C}. \end{aligned}$$

In the second inequality above we used that $\sqrt{m}2^{n-m}(\alpha_{n-m} - \alpha_{n-m-1}) = (2\beta)^{n-m}\sqrt{m} \geqq \sqrt{n}$, which follows from the inequality $(2\beta)^{n-m} = (4/3)^{n-m} \geqq \geqq \sqrt{1 + (n-m)/m}$, valid for any $n > m \geqq 2$ (if $\beta = 2/3$).

Combining the results above,

$$\begin{aligned} &\mathbf{P}\left\{\max_{0 \leqq k2^{-2m} \leqq K} |2^{-2n}T_{m,n}(k) - k2^{-2m}| \geqq \sqrt{18CKm}2^{-m} \text{ for some } n > m\right\} \\ &\quad = \mathbf{P}\left\{\max_{0 \leqq k2^{-2m} \leqq K} |T_{m,n}(k) - k2^{2(n-m)}| \geqq 3\sqrt{24CK'm}2^{2n-m-2} \right. \\ &\qquad\qquad \left. \text{for some } n > m\right\} \\ &\quad \leqq \mathbf{P}\{Z_n \geqq t_n \text{ for some } n \geqq m+1\} \\ &\quad \leqq \mathbf{P}\{Z_{m+1} \geqq t_{m+1}\} + \sum_{n=m+1}^{\infty} \mathbf{P}\{Z_n < t_n, Z_{n+1} \geqq t_{n+1}\} \\ &\quad \leqq \sum_{n=m}^{\infty} 2(K2^{2n})^{1-C} = 2(K2^{2m})^{1-C}\frac{1}{1 - 2^{2(1-C)}} \leqq 4(K2^{2m})^{1-C} \end{aligned}$$

if $C \geqq 3/2$, say. This proves (a).

The statement in (b) follows by the Borel–Cantelli lemma. □

As (67) shows, $B_n(2^{-2n}T_{m,n}(k)) = B_m(k2^{-2m})$ for any $k \geqq 0$ and for any $n > m \geqq 0$. A natural question, important particularly when looking at increments during short time intervals, is that how much time it takes for $B_n(t)$ to go from the point $B_m((k-1)2^{-2m})$ to its next "m-neighbour" $B_m(k2^{-2m})$. Is this time significantly different from 2^{-2m} for large values of m? Introducing the notation

$$\tau_{m,n}(k) = T_{m,n}(k) - T_{m,n}(k-1) \qquad (k \geqq 1,\ n > m \geqq 0), \tag{71}$$

the nth time differences of the m-neighbours are $2^{-2n}\tau_{m,n}(k)$ $(k \geqq 1)$. Note that $T_{m,n}(k) = \sum_{j=1}^{k} \tau_{m,n}(j)$, where, as can be seen from the construction and the argument below, the terms are independent and have the same distribution.

Let us look at $\tau_{m,n}(k)$ more closely. If $n = m+1$,

$$\tau_{m,m+1}(k) = T_{m+1}(k) - T_{m+1}(k-1) = \tau_{m+1}(k), \tag{72}$$

which is the double of a geometric random variable with parameter $p = 1/2$, see (14). That is, $2^{-2(m+1)}\tau_{m+1}(k)$ is the length of the time period that corresponds to the time interval $[(k-1)2^{-2m},\ k2^{-2m}]$ after the next refinement of the construction.

Similarly, each unit in $\tau_{m+1}(k)$ will bring some $\tau_{m+2}(r)$ "offsprings" after the following refinement, and so on. Hence if $n > m$ is arbitrary, then given $T_{m,n}(k-1) = j$ for some integer $j \geqq 0$, we have

$$\tau_{m,n+1}(k) = T_{n+1}(j + \tau_{m,n}(k)) - T_{n+1}(j) = \sum_{r=1}^{\tau_{m,n}(k)} \tau_{n+1}(j+r). \tag{73}$$

For given $\tau_{m,n}(k) = s$ ($s > 0$, even) its conditional distribution is the same as the distribution of a random variable T_s which is the double of a negative binomial random variable with parameters s and $p = 1/2$, described by (19) and (20). Note that this conditional distribution of $\tau_{m,n+1}(k)$ is independent of the value of $T_{m,n}(k-1)$.

Though we will not explicitly use them, it is instructive to determine some further properties of a "prototype" $\tau_{m,n} = \tau_{m,n}(1)$. A recursive formula can be given for its expectation by (73) and the full expectation formula:

$$\mu_{n+1} = \mathbf{E}(\tau_{m,n+1}) = \mathbf{E}(\mathbf{E}(\tau_{m,n+1} \mid \tau_{m,n})) = \mathbf{E}(\tau_{m,n}\mathbf{E}(\tau_{n+1}(r))) = 4\mu_n.$$

Since $\mu_{m+1} = \mathbf{E}(\tau_{m+1}(r)) = 4$, it follows that

$$\mu_n = \mathbf{E}(\tau_{m,n}) = 2^{2(n-m)}.$$

This argument also implies that

$$\mathbf{E}(2^{-2(n+1)}\tau_{m,n+1} \mid 2^{-2n}\tau_{m,n}) = 2^{-2n}\tau_{m,n}.$$

These show that the sequence $\left(2^{-2n}\tau_{m,n}\right)_{n=m+1}^{\infty}$ is a so-called *martingale*. Therefore a famous martingale convergence theorem [1, Section 5.4] implies that this sequence converges to a random variable t_m as $n\to\infty$, with probability 1, and t_m has finite expectation.

We mention that a similar recursion can be obtained for the variance that results

$$\mathbf{Var}(2^{-2n}\tau_{m,n}) < \frac{2}{3}2^{-4m}.$$

The next lemma gives an upper bound for the nth time differences of the m-neighbours by showing that during arbitrary many refinements, they cannot be "much" larger than $h=2^{-2m}$, the original time difference of the m-neighbours. More accurately, they are less than a multiple of $h^{1-\delta}$, where $\delta>0$ arbitrary.

LEMMA 6. (a) *For any $K>0$, δ such that $0<\delta<1$, and $C>2/\delta$ we have*

$$\mathbf{P}\left\{\max_{1\leqq k2^{-2m}\leqq K}|2^{-2n}\tau_{m,n}(k)-2^{-2m}|\geqq 3C2^{-2m(1-\delta)} \text{ for some } n>m\right\} \leqq \frac{K}{10}2^{-2m(\delta C-2)}.$$

(b) *For any $K>0$, and δ such that $0<\delta<1$,*

$$\max_{1\leqq k2^{-2m}\leqq K}|2^{-2n}\tau_{m,n}(k)-2^{-2m}|<\frac{7}{\delta}2^{-2m(1-\delta)},$$

with probability 1 *for all $n>m$, for all but finitely many m.*

PROOF. This proof is very similar to the proof of Lemma 5. Take any β, $1/2<\beta<1$; for example $\beta=2/3$ will do. Set $\alpha_j=1+\beta+\beta^2+\cdots+\beta^j$ for $j\geqq 0$. For any $m\geqq 0$, consider an arbitrary k, $1\leqq k\leqq K2^{2m}$. (The distribution of $\tau_{m,n}(k)$ does not depend on k.) Let

$$Z_n=|\tau_{m,n}(k)-2^{2(n-m)}|, \qquad t_n=\alpha_{n-m-1}C2^{2\delta m}2^{2(n-m)},$$

and $Y_{n+1}=|\tau_{m,n+1}(k)-4\tau_{m,n}(k)|$ for $n>m\geqq 0$. We want to apply inequality (68).

First take $n=m+1$. By (72), $\frac{1}{2}\tau_{m,m+1}(k)$ is a geometric random variable with parameter $p=1/2$. Then

$$\begin{aligned}\mathbf{P}\{Z_{m+1}\geqq t_{m+1}\} &= \mathbf{P}\left\{|\tau_{m,m+1}(k)-4|\geqq 4C2^{2\delta m}\right\}\\ &= \mathbf{P}\left\{\frac{1}{2}\tau_{m+1}(k)\geqq 2+2C2^{2\delta m}\right\}<2^{-4C\delta m},\end{aligned}$$

because of the basic property of the tail of a geometric distribution.

Second, let $n > m$ be arbitrary. By the triangle inequality, $Z_{n+1} \leqq 4Z_n + Y_{n+1}$. So we obtain

$$\begin{aligned}
&\mathbf{P}\left\{Z_n < t_n, Z_{n+1} \geqq t_{n+1}\right\} \leqq \mathbf{P}\left\{Z_n < t_n, Y_{n+1} \geqq t_{n+1} - 4t_n\right\} \\
&\leqq \sum_{s=1}^{2^{2(n-m)}+t_n} \mathbf{P}\left\{|T_s - 4s| \geqq t_{n+1} - 4t_n \mid \tau_{m,n}(k) = s\right\} \mathbf{P}\left\{\tau_{m,n}(k) = s\right\} \\
(74)\quad &\leqq \mathbf{P}\left\{\max_{1 \leqq s \leqq 2^{2(n-m)}+t_n} |T_s - 4s|/\sqrt{8} \geqq (t_{n+1} - 4t_n)/\sqrt{8}\right\} \sum_{s=1}^{\infty} \mathbf{P}\left\{\tau_{m,n}(k) = s\right\},
\end{aligned}$$

where we applied (73) and the conditional distribution of $\tau_{m,n+1}(k)$ mentioned there.

The sum in (74) is 1. Therefore we want to estimate the probability of the maximum there by using Lemma 2 with $N = 4C2^{2\delta m}2^{2(n-m)}$. This N is larger than $2^{2(n-m)} + t_n$ if m is large enough, depending on δ. (Remember that $\alpha_j < 3$ for any $j \geqq 0$ if $\beta = 2/3$.) To apply Lemma 2 we have to compare $\sqrt{2CN \log N}$ to the right-hand side of the inequality in (74):

$$\frac{\sqrt{2CN \log N}}{(t_{n+1} - 4t_n)/\sqrt{8}} = 2\left(\frac{(n-m)\log 4 + m\delta \log 4 + \log(4C)}{(4/3)^{2(n-m)}2^{2\delta m}}\right)^{1/2},$$

which is less than 1 for all $n > m$ if m is large enough, depending on δ.

Thus Lemma 2 gives that

$$\begin{aligned}
\mathbf{P}\left\{Z_n < t_n, Z_{n+1} \geqq t_{n+1}\right\} &\leqq \mathbf{P}\left\{\max_{1 \leqq s \leqq N} |T_s - 4s|/\sqrt{8} \geqq \sqrt{2CN \log N}\right\} \\
&\leqq 2N^{1-C} = 2\left(4C2^{2\delta m}2^{2(n-m)}\right)^{1-C}
\end{aligned}$$

for all $n > m$ as $m \geqq m_0(\delta, C)$.

Summing up for $n > m$, we obtain the following estimate for any given k, $1 \leqq k2^{-2m} \leqq K$, as $m \geqq m_0(\delta, C)$:

$$\begin{aligned}
&\mathbf{P}\left\{|2^{-2n}\tau_{m,n}(k) - 2^{-2m}| \geqq 3C2^{-2m(1-\delta)} \text{ for some } n > m\right\} \\
&\leqq \mathbf{P}\left\{Z_n \geqq t_n \text{ for some } n \geqq m+1\right\} \\
&\leqq \mathbf{P}\left\{Z_{m+1} \geqq t_{m+1}\right\} + \sum_{n=m+1}^{\infty} \mathbf{P}\left\{Z_n < t_n, Z_{n+1} \geqq t_{n+1}\right\} \\
&\leqq 2^{-4C\delta m} + 2^{2\delta(1-C)m}2(4C)^{1-C} \sum_{n=m+1}^{\infty} 2^{2(1-C)(n-m)} \\
&< (1/10)2^{-2m(\delta C-1)},
\end{aligned}$$

where we took into consideration that $0<\delta<1$, $C>2$, $\delta C>2$ by our assumptions.

Finally, the statement in (a) can be obtained by an application of the crude inequality (42),

$$\mathbf{P}\left\{\max_{1\leqq k\leqq K2^{2m}}|2^{-2n}\tau_{m,n}(k)-2^{-2m}|\geqq 3C2^{-2m(1-\delta)} \text{ for some } n>m\right\}$$
$$\leqq K2^{2m}\frac{1}{10}2^{-2m(\delta C-1)},$$

as $m\geqq m_0(\delta,C)$, which is equivalent to (a).

The statement in (b) follows by the Borel–Cantelli lemma with $C=$ $=7/(3\delta)$ (say). □

Now we define a certain *imbedding* of shrunk random walks $B_m(k2^{-2m})$ into the Wiener process $W(t)$.

LEMMA 7. (a) *For any $C\geqq 3/2$, $K'>K>0$, and any fixed $m\geqq m_0(C,K,K')$ there exist random time instants $t_m(k)\in[0,K']$ such that*

$$\mathbf{P}\left\{W(t_m(k))=B_m(k2^{-2m}),\ 0\leqq k2^{-2m}\leqq K\right\}\geqq 1-4(K2^{2m})^{1-C},$$

where

$$\mathbf{P}\left\{\max_{0\leqq k2^{-2m}\leqq K}|t_m(k)-k2^{-2m}|\geqq\sqrt{18CKm}2^{-m}\right\}\leqq 4(K2^{2m})^{1-C}. \tag{75}$$

Moreover, if δ is such that $0<\delta<1$, $C>2/\delta$, and $m\geqq m_1(\delta,C,K,K')$, then we also have

$$\begin{aligned}&\mathbf{P}\left\{\max_{1\leqq k2^{-2m}\leqq K}|t_m(k)-t_m(k-1)-2^{-2m}|\geqq 3C2^{-2m(1-\delta)}\right\}\\&\qquad\leqq\frac{K}{10}2^{-2m(\delta C-2)}+4(K2^{2m})^{1-C}.\end{aligned} \tag{76}$$

(b) *With probability 1, for any $K'>K>0$, $0<\delta<1$, and for all but finitely many m there exist random time instants $t_m(k)\in[0,K']$ such that*

$$W(t_m(k))=B_m(k2^{-2m})\qquad(0\leqq k2^{-2m}\leqq K),$$

where

$$\max_{0\leqq k2^{-2m}\leqq K}|t_m(k)-k2^{-2m}|\leqq\sqrt{27Km}2^{-m},$$

and

$$\max_{1\leqq k2^{-2m}\leqq K}|t_m(k)-t_m(k-1)-2^{-2m}|\leqq(7/\delta)2^{-2m(1-\delta)}.$$

PROOF. By Lemma 5(a), fixing an $m \geqq m_0(C, K, K')$, on a subset A_m of the sample space with $\mathbf{P}\{A_m\} \geqq p_m = 1 - 4(K2^{2m})^{1-C}$, one has

$$\max_{0 \leqq k2^{-2m} \leqq K} |2^{-2n}T_{m,n}(k) - k2^{-2m}| < \sqrt{18CKm}\,2^{-m}, \tag{77}$$

for each $n > m$. In particular, the time instants $2^{-2n}T_{m,n}(k)$ are bounded from below by 0 and from above by $K + \sqrt{18CKm}\,2^{-m} \leqq K'$. (Assume that $m_0(C, K, K')$ is chosen so.)

Applying a truncation $t^*_{m,n}(k) = \min\{K', 2^{-2n}T_{m,n}(k)\}$, for each k, $0 \leqq k2^{-2m} \leqq K$, we get a sequence in n bounded over the whole sample space, equal to the original one for $\omega \in A_m$. It follows from the classical Weierstrass theorem [7, Section 2.42], that every bounded sequence of real numbers contains a convergent subsequence. To be definite, let us take the lower limit [7, Section 3.16] of the sequence:

$$t_m(k) = \liminf_{n \to \infty} t^*_{m,n}(k). \tag{78}$$

Then $t_m(k) \in [0, K']$.

By Theorem 3, with probability 1 the sample-functions of $B_n(t)$ uniformly converge to the corresponding sample-functions of the Wiener process that are uniformly continuous on $[0, K']$. (A continuous function on a closed interval is uniformly continuous [7, Section 4.19].) Thus (67) implies that for each k, $0 \leqq k2^{-2m} \leqq K$, we have $W(t_m(k)) = B_m(k2^{-2m})$, with probability at least p_m (on the set A_m where the truncated sequences coincide with the original ones).

To show it in detail, take any $\epsilon > 0$, any k $(0 \leqq k2^{-2m} \leqq K)$, and a subsequence $t^*_{m,n_i}(k)$ converging to $t_m(k)$ as $i \to \infty$. Then

$$\begin{aligned}
&|W(t_m(k)) - B_m(k2^{-2m})| = |W(t_m(k)) - B_{n_i}(2^{-2n_i}T_{m,n_i}(k))| \\
&\leqq |W(t_m(k)) - W(2^{-2n_i}T_{m,n_i}(k))| + |W(2^{-2n_i}T_{m,n_i}(k)) - B_{n_i}(2^{-2n_i}T_{m,n_i}(k))| \\
&< \epsilon/2 + \epsilon/2 = \epsilon,
\end{aligned}$$

where the last inequality holds on the set A_m, for all but finitely many n_i. Since ϵ was arbitrary, it follows that $|W(t_m(k)) - B_m(k2^{-2m})| = 0$ on A_m.

Further, taking a limit in (77) with $n = n_i$ as $i \to \infty$ (on the set A_m), one obtains (75). Also, taking a similar limit in Lemma 6(a), $2^{-2n}\tau_{m,n_i}(k) \to t_m(k) - t_m(k-1)$ on the set A_m, and (76) follows.

The statements in (b) can be obtained similarly as in (a), applying Lemmas 5(b) and 6(b), or from (a) by the Borel–Cantelli lemma. □

We mention that for any $k \geqq 0$ and $m \geqq 0$, the sequence $2^{-2n}T_{m,n}(k)$ in fact converges to $t_m(k)$ with probability 1 as $n \to \infty$. However, a "natural" proof of this fact requires the martingale convergence theorem mentioned above, before Lemma 6, a tool of more advanced nature than the ones we use in this paper.

Next we want to show that the random time instants $s_m(k)$ of the Skorohod imbedding (65) and the $t_m(k)$'s defined in (78) are essentially the same. This requires a recollection of some properties of random walks.

We want to estimate the probability that with given positive integers j, x, u and r a random walk S_i goes from a point $|S_j| = x$ to $|S_{j+k}| = x + y$ so that $|S_{j+i}| < x + y$ while $1 \leqq i < k$ for some $y \leqq r$ and $k \geqq u$, where k, y, and i are also positive integers.

The first passage distribution given in [2, Section III,7] can be applied here:

$$\mathbf{P}\left\{S_0 = 0, S_i < y \ (1 \leqq i < k), S_k = y\right\} = \frac{y}{k}\binom{k}{(k+y)/2} 2^{-k}.$$

Hence, by Theorem 1,

$$\begin{aligned}
\mathbf{P}&\left\{|S_j| = x, |S_{j+i}| < x + y \ (1 \leqq i < k), |S_{j+k}| = x + y \ \text{ for some } \ y \leqq r\right\} \\
&\leqq \sum_{y=1}^{r} \frac{y}{k}\binom{k}{(k+y)/2} 2^{-k} \\
&\leqq (1+\epsilon)\frac{r}{k}\left(\Phi(r/\sqrt{k}) - \Phi(0)\right) \\
&\leqq (1+\epsilon)\frac{r}{k}\frac{r}{\sqrt{k}}\frac{1}{\sqrt{2\pi}} \leqq \frac{r^2}{k^{3/2}},
\end{aligned}$$

where $\epsilon > 0$ is arbitrary, say equals 1, and $k \geqq k_0$.

So the larger the value of k is, the smaller estimate of the probability we get. Thus for all positive integers j, x, r, and $u \geqq k_0$,

$$\begin{aligned}
&\mathbf{P}\left\{|S_j| = x, |S_{j+i}| < x + y \ (1 \leqq i < k),\right. \\
&\left.|S_{j+k}| = x + y \ \text{ for some } \ y \leqq r, \ k \geqq u\right\} \leqq r^2/u^{3/2},
\end{aligned} \tag{79}$$

independently of the values of j and x.

THEOREM 4. *The stopping times $s_m(k)$ $(k \geqq 0)$ of the Skorohod imbedding are equal to the time instants $t_m(k)$ of the imbedding defined in Lemma 7 on the set A_m of the sample space given by* (69), *with the possible exception of a zero probability subset.*

Therefore all statements in Lemma 7 hold when $s_m(k)$ replaces $t_m(k)$.

PROOF. Fix an $m \geqq m_0(C, K, K')$, where $m_0(C, K, K')$ is the same as in Lemma 7. Let the subset A_m of the sample space be given by (69).

Take $k = 1$ first. Since $s_m(1)$ is the smallest time instant where $|W(t)|$ is equal to 2^{-m}, and $|W(t_m(1))| = 2^{-m}$ on the set A_m, it follows that $s_m(1) \leqq$ $\leqq t_m(1)$ on A_m. We want to show that on A_m the event $\{s_m(1) < t_m(1)\}$ has zero probability.

Indirectly, let us suppose that $\delta_m = t_m(1) - s_m(1) > 0$ on a subset C_m of A_m with positive probability. By (67), the first time instant where $|B_n(t)|$ equals $|B_m(2^{-2m})| = 2^{-m}$ is $2^{-2n}T_{m,n}(1)$ $(n > m)$. So $|B_n(t)| < 2^{-m}$ if $0 \leqq t < 2^{-2n}T_{m,n}(1)$. On the other hand, by (55), $2^{-m} - n2^{-n/2} \leqq |B_n(s_m(1))| < 2^{-m}$ for $n \geqq N_1(\omega)$ on a probability 1 ω-set. (Remember that $|W(s_m(1))| = 2^{-m}$.)

Since $\delta_m > 0$ on the set C_m, there exists an $N_2(\omega)$ such that $n2^{-n/2} < \delta_m/2$ for $n \geqq N_2(\omega)$.

By (78), $t_m(1) = \liminf_{n\to\infty} 2^{-2n}T_{m,n}(1)$ on the set A_m. The properties of the lower limit [7, Section 3.17] imply that on the subset C_m there exists an $N_3(\omega)$ such that $2^{-2n}T_{m,n}(1) > t_m(1) - \delta_m/2$ for $n \geqq N_3(\omega)$.

Set $N(\omega) = \max\{N_1(\omega), N_2(\omega), N_3(\omega)\}$ for $\omega \in C_m$. Since $B_n(t) = 2^{-n}\tilde{S}_n(t2^{2n})$, the statements above imply that on the set C_m the random walk $\tilde{S}_n(t)$ have the following properties for $n \geqq N(\omega)$:

(a) $|\tilde{S}_n(s_m(1)2^{2n})| \geqq 2^{n-m} - n2^{n/2}$,

(b) $|\tilde{S}_n(t)| < 2^{n-m}$ for $s_m(1)2^{2n} \leqq t < T_{m,n}(1)$, where $T_{m,n}(1) - s_m(1) > (\delta_m/2)2^{2n} > n2^{3n/2}$,

(c) $|\tilde{S}_n(T_{m,n}(1))| = 2^{n-m}$.

Let $D_{m,n}$ denote the subset of C_m on which (a), (b), and (c) hold for a fixed n. Since $D_{m,n} \subset D_{m,n+1}$ for each n, by the continuity property of probability [7, Section 11.3], we have $\lim_{n\to\infty} \mathbf{P}\{D_{m,n}\} = \mathbf{P}\{C_m\} > 0$. This implies that there exists an integer n_0 such that $\mathbf{P}\{D_{m,n}\} \geqq \frac{1}{2}\mathbf{P}\{C_m\} > 0$ holds for all $n \geqq n_0$ (say). In other words, for all large enough values of n, the probability of the event that (a), (b), and (c) hold simultaneously is larger than a fixed positive number.

To get a contradiction, we apply (79) to $\tilde{S}_n(t)$, with $r = n2^{n/2}$ and $u = n2^{3n/2}$. Theorem 1, that was used to deduce (79), still applies since $r = o(u^{2/3})$, i.e. $r/\sqrt{u} = o(u^{1/6})$. Now the first passage time when $|\tilde{S}_n(t)|$ hits 2^{-2m} is $T_{m,n}(1)$. Thus the probability that $\tilde{S}_n(t)$ satisfies (a), (b), and (c) simultaneously is less than or equal to

$$\frac{r^2}{u^{3/2}} = \frac{(n2^{n/2})^2}{(n2^{3n/2})^{3/2}} = \frac{\sqrt{n}}{2^{5n/4}},$$

which goes to zero as $n \to \infty$. This contradicts the statement above that for all large enough value of n, the event that (a), (b), and (c) hold has a probability larger than a fixed positive number. This proves the lemma for $k = 1$: $s_m(1) = t_m(1)$ on the set A_m, with the possible exception of a zero probability subset.

For $k>1$, one can proceed by induction. Assume that $s_m(k-1)=t_m(k-1)$ holds on A_m except possibly for a subset of probability zero. The proof that then $s_m(k) = t_m(k)$ holds as well is essentially the same as the proof of the case $k=1$ above. It is true because on one hand $s_m(k)$ is defined recursively in (65), using $s_m(k-1)$, the same way as $s_m(1)$ is defined. On the other hand, by (71), $T_{m,n}(k)=T_{m,n}(k-1)+\tau_{m,n}(k)$, where the $\tau_{m,n}(k)$ is defined the same way as $\tau_{m,n}(1)=T_{m,n}(1)$. Also, remember that on the set A_m, $t_m(j)=\liminf_{n\to\infty} T_{m,n}(j)$ for $j=k-1$ and $j=k$. □

6. Some properties of the Wiener process

Theorem 3 above indicates that the sample-functions of the Wiener process are arbitrarily close to the sample-functions of $B_n(t)$ if n is large enough, with probability 1. The sample-functions of $B_n(t)$ are broken lines that have a chance of 1/2 to turn and have a corner at any multiple of time $1/2^{2n}$, so at more and more instants of time as $n\to\infty$. Moreover, the magnitude of the slopes of the line segments that make up the graph of $B_n(t)$ is

$$\frac{1/2^n}{1/2^{2n}}=2^n\to\infty \text{ as } n\to\infty.$$

Therefore one would suspect that the sample-functions of the Wiener process are typically nowhere differentiable. As we will see below, this is really true. Thus typical sample-functions of the Wiener process belong to the "strange" class of the everywhere continuous but nowhere differentiable functions.

THEOREM 5. *With probability* 1, *the sample-functions of the Wiener process are nowhere differentiable.*

PROOF. It suffices to show that with probability 1, the sample-functions are nowhere differentiable on any interval $[0,K]$. Put $K_0=(3/2)K>0$ (say). Then with probability 1, for all sample-functions and for all but finitely many m there exist time instants $t_m(k)$ $(0\leqq k2^{-2m}\leqq K_0)$ with the properties described in Lemma 7(b). In particular,

$$\max_{0\leqq k2^{-2m}\leqq K_0} t_m(k)\geqq K_0-\sqrt{27K_0 m}2^{-m}>K$$

if m is large enough.

Fix an ω in this probability 1 subset of the sample space. This defines a specific sample-function of $W(t)$ and specific values of the random time instants $t_m(k)$. (To simplify the notation, in this proof we suppress the argument ω.) Then choosing an arbitrary point $t\in[0,K]$, for each m large enough, one has $t_m(k-1)\leqq t<t_m(k)$ for some k, $0<k2^{-2m}\leqq K_0$. Taking for instance $\delta=1/4$ in Lemma 7(b), we get $t_m(k)-t_m(k-1)\leqq 29\cdot 2^{-(3/2)m}$ and

$$|W(t_m(k))-W(t_m(k-1))|=|B_m(k2^{-2m})-B_m((k-1)2^{-2m})|=2^{-m}.$$

Set $t_m^* = t_m(k)$ if $|W(t) - W(t_m(k))| \geqq |W(t) - W(t_m(k-1))|$ and $t_m^* = t_m(k-1)$ otherwise. Then $|W(t) - W(t_m^*)| \geqq (1/2)2^{-m}$. So $|t_m^* - t| \leqq 29 \cdot 2^{-(3/2)m} \to 0$ and

$$\left| \frac{W(t_m^*) - W(t)}{t_m^* - t} \right| \geqq \frac{(1/2)2^{-m}}{29 \cdot 2^{-(3/2)m}} = \frac{1}{58} 2^{m/2} \to \infty,$$

as $m \to \infty$. This shows that the given sample-function cannot be differentiable at any point $t \in [0, K]$. □

It has important consequences in the definition of stochastic integrals that, as shown below, the graph of a typical sample-function of the Wiener process has infinite length. In general, (the graph of) a function f defined on an interval $[a, b]$ has *finite length* (or f is said to be of *bounded variation* on $[a, b]$) if there exists a finite constant c such that for any *partition* $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$, the sum of the absolute values of the corresponding changes does not exceed c:

$$\sum_{j=1}^{n} |f(x_j) - f(x_{j-1})| \leqq c.$$

The smallest c with this property is called the *total variation* of f over $[a, b]$, denoted $V(f(t),\ a \leqq t \leqq b)$. Otherwise we say that the graph has *infinite length*, or f is of *unbounded variation* on $[a, b]$.

First let us calculate the total variation of a sample-function of $B_m(t)$ over an interval $[0, K]$. Each sample-function of $B_m(t)$ over $[0, K]$ is a broken line that consists of $K2^{2m}$ line segments with changes of magnitude 2^{-m}. So for any sample-function of $B_m(t)$,

$$V(B_m(t),\ 0 \leqq t \leqq K) = K2^{2m}2^{-m} = K2^m, \tag{80}$$

which tends to infinity as $m \to \infty$.

LEMMA 8. *For any $K' > 0$, the sample-functions of the Wiener process over $[0, K']$ have infinite length (i.e. are of unbounded variation) with probability* 1.

PROOF. By Lemma 7, for any $C \geqq 3/2$, $K' > K > 0$, and $m \geqq m_0(C, K, K')$ there exist time instants $t_m(k) \in [0, K']$ such that

$$\mathbf{P}\left\{W(t_m(k)) = B_m(k2^{-2m}),\ 0 \leqq k2^{-2m} \leqq K\right\} \geqq 1 - 4(K2^{2m})^{1-C}. \tag{81}$$

For each $m \geqq 0$ define the following event:

$$C_m = \left\{V(W(t),\ 0 \leqq t \leqq K') < K2^m\right\}.$$

Then $C_m \subset C_{m+1}$ for any $m \geqq 0$.

For any sample-function of $W(t)$, take the partition $0 = t_m(0) < t_m(1) < \dots < t_m(K2^{2m})$. (To alleviate the notation, we suppress the dependence on ω.) By (81), for any $m \geqq m_0(C, K, K')$, the sum of the corresponding absolute changes is equal to $K2^{2m}2^{-m} = K2^m$, with probability at least $1 - 4(K2^{2m})^{1-C}$.

This shows that then $\mathbf{P}\{C_m\} < 4(K2^{2m})^{1-C}$. Take the event

$$C_\infty = \{V(W(t),\ 0 \leqq t \leqq K') < \infty\}.$$

The continuity property of probability implies that $\mathbf{P}\{C_m\} \to \mathbf{P}\{C_\infty\}$ as $m \to \infty$, that is, $\mathbf{P}\{C_\infty\} = 0$. □

The next lemma shows a certain uniform continuity property of the Wiener process. An interesting consequence of the lemma is that for any $u > 0$ the probability that $|W(t) - W(s)| \geqq u$ holds for some $s, t \in [0, K]$, $|t - s| \leqq h$ can be made arbitrarily small if a small enough h is chosen. More accurately, the lemma shows that only with small probability can the increment of the Wiener process be larger than $c\sqrt{h}$ if the constant c is large enough. Now $\sqrt{h}$ is much larger than h for small values of h, so this also indicates why sample-functions of the Wiener process are not differentiable. At the same time it gives a rough measure of the so-called *modulus of continuity* of the process. Basically, the proof relies on Theorem 1a and Theorem 3.

LEMMA 9. *For any $K>0$, $0<\delta<1$, and $u>0$ there exists an $h_0(K, \delta, u)>0$ such that*

$$\mathbf{P}\left\{\max_{s,t\in[0,K],\ |t-s|\leqq h} |W(t) - W(s)| \geqq u\right\} \leqq 7e^{-\frac{u^2}{2h}(1-\delta)}, \tag{82}$$

for all positive $h \leqq h_0(K, \delta, u)$.

PROOF. First we choose a large enough $C \geqq 3/2$ such that $2/(C-1) < \delta/2$. For instance, $C = 1 + (6/\delta)$ will do.

By (54), the probability in (82) cannot exceed

$$6(K2^{2n})^{-6/\delta} + \mathbf{P}\left\{\max_{0\leqq s\leqq K-h}\ \max_{s\leqq t\leqq s+h} |B_n(t) - B_n(s)| \geqq u - 2n2^{-n/2}\right\}, \tag{83}$$

for $n \geqq n_0(K, \delta)$. (Remember that $1 - C = -6/\delta$ now.)

By definition, $B_n(t) = 2^{-n}\tilde{S}_n(t2^{2n})$ for $t \geqq 0$. For each $s \leqq t$ from $[0, K]$ and $n \geqq n_0(K, \delta)$ take the integers $s_n = \lceil s2^{2n} \rceil$ and $t_n = \max\{s_n, \lfloor t2^{2n} \rfloor\}$. ($\lceil x \rceil$ denotes the smallest integer greater than or equal to x, while $\lfloor x \rfloor$ is the largest integer smaller than or equal to x.)

Then $|t_n - t2^{2n}| \leqq 1$ and so $|\tilde{S}_n(t_n) - \tilde{S}_n(t2^{2n})| \leqq 1$, similarly for s_n. Moreover, $0 \leqq t_n - s_n \leqq h2^{2n}$ if $0 \leqq t - s \leqq h$. Hence (83) does not exceed

$$\begin{aligned} &6(K2^{2n})^{-6/\delta} + \\ &+\mathbf{P}\left\{\max_{0\leqq j\leqq K2^{2n}}\ \max_{0\leqq k\leqq h2^{2n}} |\tilde{S}_n(j+k) - \tilde{S}_n(j)| \geqq 2^n(u - 2n2^{-n/2}) - 2\right\}, \end{aligned} \tag{84}$$

for $n \geqq n_0(K, \delta)$.

The distribution of $\tilde{S}_n(j+k) - \tilde{S}_n(j)$ above is the same as the distribution of a random walk $S(k)$, for any value of $k \geqq 0$, independently of $j \geqq 0$. Also, the largest possible value of $|S(k)|$ is k. Therefore by Theorem 1a, the inequality (5), and the crude estimate (42),

$$\mathbf{P}\left\{\max_{0 \leqq k \leqq h2^{2n}} |\tilde{S}_n(j+k) - \tilde{S}_n(j)| \geqq 2^n(u - 2n2^{-n/2} - 2 \cdot 2^{-n})\right\}$$
$$\leqq \mathbf{P}\left\{\max_{u\sqrt{1-\delta/2}2^n \leqq k \leqq h2^{2n}} \frac{|S(k)|}{\sqrt{k}} \geqq \frac{u}{\sqrt{h}}\sqrt{1-\delta/2}\right\} \leqq h2^{2n}e^{-\frac{u^2}{2h}(1-\delta/2)}.$$

Here it was assumed that $2n2^{-n/2} + 2 \cdot 2^{-n} \leqq u(1 - \sqrt{1-\delta/2})$, which certainly holds if $n \geqq n_1(K, \delta, u) \geqq n_0(K, \delta)$. Also, we assumed that $\frac{u}{\sqrt{h}}\sqrt{1-\delta/2} \geqq$ $\geqq 3/\sqrt{2\pi}$, see (6), which is true if h is small enough, depending on δ and u.

Consequently, applying the crude estimate (42) again for (84), we obtain

$$\mathbf{P}\left\{\max_{s,t \in [0,K],\ |t-s| \leqq h} |W(t) - W(s)| \geqq u\right\}$$
$$\leqq 6(K2^{2n})^{-6/\delta} + K2^{2n}h2^{2n}e^{-\frac{u^2}{2h}(1-\delta/2)}$$
$$= 6e^{-\frac{6}{\delta}(\log K + 2n\log 2)} + Khe^{4n\log 2 - \frac{u^2}{2h}(1-\delta/2)}.$$

Now we select an integer $n \geqq n_1(K, \delta, u)$ such that $-\frac{6}{\delta}(\log K + 2n\log 2) \leqq$ $\leqq -\frac{u^2}{2h}$. The choice

$$n = \left\lceil \frac{1}{2\log 2}\left(\frac{u^2}{2h}\frac{\delta}{6} - \log K\right)\right\rceil$$

will do if h is small enough, $0 < h \leqq h_0(K, \delta, u)$, so that $n \geqq n_1(K, \delta, u) \geqq 2$. Then $n \leqq \frac{3}{2}\frac{1}{2\log 2}\left(\frac{u^2}{2h}\frac{\delta}{6} - \log K\right)$ holds as well.

With this n we have $4n\log 2 \leqq \frac{u^2}{2h}\delta/2 + \log(K^{-3})$, and so

$$\mathbf{P}\left\{\max_{s,t \in [0,K],\ |t-s| \leqq h} |W(t) - W(s)| \geqq u\right\}$$
$$\leqq 6e^{-\frac{u^2}{2h}} + KhK^{-3}e^{-\frac{u^2}{2h}(1-\delta/2-\delta/2)}$$
$$\leqq (6 + h/K^2)e^{-\frac{u^2}{2h}(1-\delta)}.$$

If $K \geqq 1$, then $h/K^2 \leqq 1$ and (82) follows. If $K < 1$, the maximum in (82) cannot exceed the maximum over the interval $[0,1]$. Then taking $h_0(K, \delta, u) = h_0(1, \delta, u)$, (82) follows again. $\square$

7. A preview of stochastic integrals

To show how stochastic integrals come as natural tools when working with differential equations including random effects, and what kind of problems arise when one wants to define them, let us start with the simplest ordinary differential equation

$$x'(t) = f(t) \qquad (t \geqq 0),$$

where f is a continuous function. If $x(0)$ is given, its unique solution can be obtained by integration,

$$x(t) - x(0) = \int_0^t f(s)\, ds \qquad (t \geqq 0).$$

Now we modify this simple model by introducing a random term, very customary in several applications:

$$x'(t) = f(t) + g(t)W'(t) \qquad (t \geqq 0),$$

where f and g are continuous random functions and $W'(t)$ is the so-called *white noise* process. Now we know from Theorem 5 that $W'(t)$ does not exist (at least not in the ordinary sense), but after integration we may get some meaningful solution,

$$x(t) - x(0) = \int_0^t f(s) ds + \int_0^t g(s) dW(s) \qquad (t \geqq 0).$$

The second integral here is what one wants to call a stochastic integral if it can be defined properly.

A natural idea to define such a stochastic integral is to define it as a *Riemann–Stieltjes integral* [7, Chapter 6] for each sample-function separately. It means that one takes partitions $0 = s_0 < s_1 < \cdots < s_{n-1} < s_n = t$, and Riemann–Stieltjes sums

$$\sum_{k=1}^{n} g(u_k)(W(s_k) - W(s_{k-1})),$$

where $u_k \in [s_{k-1}, s_k]$ is arbitrary. (We suppress the argument ω that would refer to a specific sample-function in order to alleviate the notation.) Then one would hope that as the norm of the partition $\|\mathcal{P}\| = \max_{1 \leqq k \leqq n} |s_k - s_{k-1}|$ tends to 0, the Riemann–Stieltjes sums converge to the same limit when fixing a specific point ω in the sample space.

One problem is that it cannot happen to all continuous random functions g. The reason is that $W(s)$ has unbounded variation over the interval $[0,t]$ —as we saw it in Lemma 8. The random function g could be chosen so that a Riemann–Stieltjes sum gets arbitrary close to the total variation, which is ∞. Naturally, this is the case with not only the Wiener process, but with any process whose sample functions have unbounded variation, see e.g. [5, Section 4.7].

But there is another problem connected to the choice of the points $u_k \in \in [s_{k-1}, s_k]$ in the Riemann–Stieltjes sums above. This choice unfortunately does matter, not like in the case of ordinary integration. The reason is again the unbounded variation of the sample-functions. The easiest way to illustrate it is using *discrete stochastic integrals*, that is, sums of random variables. (Such a sum is essentially the same as a Riemann–Stieltjes sum above.)

So let $S_0 = 0$, $S_n = \sum_{k=1}^{n} X_k$ is a (simple, symmetric) random walk, just like in Section 1. In the following examples S_n will play the role of the function $g(t)$ above, and the white noise process $W'(t)$ is substituted by the increments X_n. In the first case (that corresponds to an *Itô-type stochastic integral*), we define the discrete stochastic integral as $\sum_{k=1}^{n} S_{k-1} X_k$. Observe that in this case the integrand is always taken at the left endpoint of the subintervals. A usual reasoning behind this is that X_k gives the "new information" in each term, while the integrand S_{k-1} depends only on the past, that is, *non-anticipating*: independent of the future values $X_k, X_{k+1}, \ldots$.

This discrete stochastic integral can be evaluated explicitly as

$$\sum_{k=1}^{n} S_{k-1} X_k = \sum_{k=1}^{n} S_{k-1} (S_k - S_{k-1})$$
$$= \frac{1}{2} \sum_{k=1}^{n} (S_k^2 - S_{k-1}^2) - \frac{1}{2} \sum_{k=1}^{n} (S_k - S_{k-1})^2 = \frac{S_n^2}{2} - \frac{n}{2}.$$

Here we used that the first resulting sum telescopes and $S_0^2 = 0$, while each term $(S_k - S_{k-1})^2$ in the second resulting sum is equal to 1. The interesting feature of the result is that it contains the non-classical term $-n/2$. The "non-classical" phrase refers to the fact that $\int_0^{s_n} s\,ds = {s_n}^2/2$. Altogether, this formula is a special case of the important *Itô formula*, one of our main subjects from now on.

Of course, it is also interesting to see what happens if the integrand is always evaluated at the right endpoints of the subintervals:

$$\sum_{k=1}^{n} S_k X_k = \sum_{k=1}^{n} S_k (S_k - S_{k-1})$$
$$= \frac{1}{2} \sum_{k=1}^{n} (S_k^2 - S_{k-1}^2) + \frac{1}{2} \sum_{k=1}^{n} (S_k - S_{k-1})^2 = \frac{S_n^2}{2} + \frac{n}{2}.$$

Note that the non-classical term is $+n/2$ here.

Taking the arithmetical average of the two formulae above we obtain a *Stratonovich-type stochastic integral*, which does not contain a non-classical term:

$$\sum_{k=1}^{n} \frac{S_{k-1}+S_k}{2} X_k = \sum_{k=1}^{n} S(k-\frac{1}{2}) X_k = \frac{S_n^2}{2}.$$

On the other hand, this type of integral has other disadvantages compared to the Itô-type one, resulting from the fact that here the integrand is "anticipating", not independent of the future.

After showing these (and other) examples in a seminar, P. Révész asked the question if there is a general method to evaluate discrete stochastic integrals of the type $\sum_{k=1}^{n} f(S_{k-1})X_k$ in closed form, where f is a given function defined on the set of integers $\mathbf{Z}$. In other words, does there exist a discrete Itô formula in general? The answer is yes, and fortunately it is quite elementary to see.

But before turning to this, let us see the relationship of such a formula to an alternative way of defining certain stochastic integrals. This important type of stochastic integrals is $\int_0^K f(W(s))dW(s)$, where $K > 0$ and f is a continuously differentiable function. In other words, the integrand is a smooth function of the Wiener process. The traditional definition of the Itô-type integral in this case goes quite similarly to the Riemann–Stieltjes integral.

Take an arbitrary partition $\mathcal{P} = \{0 = s_0, s_1, \dots, s_{n-1}, s_n = K\}$ on the time axis, and a corresponding Riemann-Stieltjes sum, evaluating the function always at the left endpoints of the subintervals,

$$\sum_{k=1}^{n} f(W(s_{k-1}))(W(s_k) - W(s_{k-1})).$$

This sum is a random variable, corresponding to the given partition. It can be proved that these random variables converge e.g. *in probability* to a certain random variable I, as the norm of the partition goes to 0. This random variable I is then called the Itô integral. We mention that "in probability" convergence means that for any $\epsilon > 0$ there exists a $\delta > 0$ such that

$$\mathbf{P}\left\{|I - \sum_{k=1}^{n} f(W(s_{k-1}))(W(s_k) - W(s_{k-1}))| \geq \epsilon\right\} < \epsilon,$$

as $||\mathcal{P}|| < \delta$.

The alternative method that we will follow in this paper is better suited to the relationship between the Wiener process and random walks discussed above. Mathematically, it somewhat reminds a *Lebesgue–Stieltjes integral* [7, Chapter 11]. The idea is that we first take a dyadic partition on the spatial axis, each subinterval having the length 2^{-m}, where m is a non-negative integer. Then we determine the corresponding first passage times $s_m(1), s_m(2), \ldots$ of the Skorohod imbedding as explained above. These time instants can be considered as a random partition on the time axis that in general depends on the considered sample-function.

By Lemma 7b and Theorem 4, with probability 1, for any $K' > 0$ and for all but finitely many m, each $s_m(k)$ lies in the interval $[0, K']$ and $W(s_m(k)) = B_m(k2^{-2m})$, $0 \leq k2^{-2m} \leq K$. The shrunk random walk $B_m(t)$ can be expressed in terms of ordinary random walks by (40) as $B_m(k2^{-2m}) = 2^{-m}\tilde{S}_m(k)$. Now our definition of the Itô integral will be

$$\lim_{m\to\infty} \sum_{k=1}^{K2^{2m}} f(W(s_m(k-1)))\,(W(s_m(k)) - W(s_m(k-1))). \tag{85}$$

We will show later that this sum, which can be evaluated for each sample-function separately, converges with probability 1. Our method will be to find an other form of this sum by a discrete Itô formula and to apply the limit to the equivalent form so obtained.

8. A discrete Itô formula

Let f be a function defined on the set of integers $\mathbf{Z}$. First we define *trapezoidal sums* of f by

$$T_{j=0}^{k}\, f(j) = \epsilon_k \left\{ \frac{1}{2} f(0) + \sum_{j=1}^{|k|-1} f(\epsilon_k j) + \frac{1}{2} f(k) \right\}, \tag{86}$$

where $k \in \mathbf{Z}$ (so k can be negative as well!) and

$$\epsilon_k = \begin{cases} 1 & \text{if } k > 0 \\ 0 & \text{if } k = 0 \\ -1 & \text{if } k < 0. \end{cases} \tag{87}$$

The reason behind the -1 factor when $k < 0$ is the analogy with integration: when the upper limit of the integration is less than the lower limit, one can exchange them upon multiplying the integral by -1.

The next statement that we will call a *discrete Itô formula* is a purely algebraic one. It is reflected by the fact that though we will apply it exclusively for random walks, the lemma holds for any numerical sequence $X_r = \pm 1$, irrespective of any probability assigned to them.

LEMMA 10. *Take any function f defined on $\mathbf{Z}$, any sequence $X_r = \pm 1$ ($r \geqq 1$), and let $S_0 = 0$, $S_n = X_1 + X_2 + \cdots + X_n$ ($n \geqq 1$). Then the following statements hold:*

DISCRETE ITÔ FORMULA

$$T_{j=0}^{S_n} f(j) = \sum_{r=1}^{n} f(S_{r-1})X_r + \frac{1}{2}\sum_{r=1}^{n} \frac{f(S_r) - f(S_{r-1})}{X_r},$$

and

DISCRETE STRATONOVICH FORMULA

$$T_{j=0}^{S_n} f(j) = \sum_{r=1}^{n} \frac{f(S_{r-1}) + f(S_r)}{2} X_r .$$

PROOF. By the definition of a trapezoidal sum,

$$T_{j=0}^{S_r} f(j) - T_{j=0}^{S_{r-1}} f(j) = X_r \frac{f(S_{r-1}) + f(S_r)}{2}, \tag{88}$$

since if $S_r - S_{r-1} = X_r$ equals 1, one has to add a term $(f(S_{r-1}) + f(S_r))/2$, while if $X_r = -1$, one has to subtract this term.

Since $X_r = \pm 1$, the right-hand side of (88) can be written as

$$T_{j=0}^{S_r} f(j) - T_{j=0}^{S_{r-1}} f(j) = f(S_{r-1})X_r + \frac{1}{2}\frac{f(S_r) - f(S_{r-1})}{X_r}. \tag{89}$$

By summing up (89), respectively (88), for $r = 1, 2, \ldots, n$ we obtain the statements of the lemma, since the sum telescopes and $T_{j=0}^{S_0} f(j) = 0$:

$$\sum_{r=1}^{n} \left(T_{j=0}^{S_r} f(j) - T_{j=0}^{S_{r-1}} f(j) \right) = T_{j=0}^{S_n} f(j). \qquad \square$$

We need a version of Lemma 10 that can be applied for shrunk random walks $B_m(t)$ as well. Therefore we define trapezoidal sums of a function f over an equidistant partition with points $x = j\Delta x$, where $\Delta x > 0$ and j changes over the set of integers $\mathbf{Z}$. Here the function f is assumed to be defined on the set of real numbers $\mathbf{R}$. So a corresponding trapezoidal sum is

$$T_{x=0}^{a} f(x)\Delta x = \epsilon_a \Delta x \left\{ \frac{1}{2} f(0) + \sum_{j=1}^{(|a|/\Delta x)-1} f(\epsilon_a j \Delta x) + \frac{1}{2} f(a) \right\}, \tag{90}$$

where a is assumed to be an integer multiple of Δx and ϵ_a is defined according to (87). In the sequel this definition will be applied with $\Delta x = 2^{-m}$. We write the corresponding version of Lemma 10 directly for shrunk random walks $B_m(t)$, though this lemma is of purely algebraic nature as well.

LEMMA 11. *Take any function f defined on $\mathbf{R}$, any real $K > 0$, and fix a non-negative integer m. Consider shrunk random walks $B_m(r2^{-2m}) = 2^{-m}\tilde{S}_m(r)$ $(r \geqq 0)$. Then the following statements hold ($\Delta x = 2^{-m}$, $\Delta t = 2^{-2m}$):*

ITÔ CASE

$$T_{x=0}^{B_m(K_m)} f(x)\Delta x = \sum_{r=1}^{\lfloor K/\Delta t \rfloor} f(B_m((r-1)\Delta t))(B_m(r\Delta t) - B_m((r-1)\Delta t))$$

$$+\frac{1}{2}\sum_{r=1}^{\lfloor K/\Delta t \rfloor} \frac{f(B_m(r\Delta t)) - f(B_m((r-1)\Delta t))}{B_m(r\Delta t) - B_m((r-1)\Delta t)}\Delta t, \tag{91}$$

and

STRATONOVICH CASE

$$T_{x=0}^{B_m(K_m)} f(x)\Delta x$$

$$= \sum_{r=1}^{\lfloor K/\Delta t \rfloor} \frac{f(B_m((r-1)\Delta t)) + f(B_m(r\Delta t))}{2}(B_m(r\Delta t) - B_m((r-1)\Delta t)), \tag{92}$$

where $K_m = \lfloor K/\Delta t \rfloor \Delta t$.

PROOF. The proof is essentially the same as in case of Lemma 10, therefore omitted. □

Now recall Lemma 7b and Theorem 4. With probability 1, for any $K' > K$ and for all but finitely many m there exist random time instants $s_m(r) \in [0, K']$ (the first passage times of the Skorohod imbedding) such that $W(s_m(r)) = B_m(r\Delta t)$ and

$$\max_{0 \leqq r\Delta t \leqq K} |s_m(r) - r\Delta t| \leqq \sqrt{27Km}\,2^{-m}, \tag{93}$$

going to 0 as $m \to \infty$.

In this light the shrunk random walks $B_m(t)$ can be replaced by the Wiener process in (91) and (92). Then the first sum on the right-hand side of (91) becomes exactly the one whose limit as $m \to \infty$ is going to be our definition of Itô integral by (85). Similarly, the right-hand side of (92) is the one whose limit will be our definition of the Stratonovich integral.

The most important feature of Lemma 11 is that these limits can be evaluated in terms of limits of other, simpler sums. An other gain is that after performing the limits, we will immediately obtain the important Itô and Stratonovich formulae for the corresponding types of stochastic integrals.

9. Stochastic integrals and the Itô formula

THEOREM 6. *Let f be a continuously differentiable function on the set of real numbers* $\mathbf{R}$, *and $K > 0$. For $m \geqq 0$ and $k \geqq 0$ take the first passage times $s_m(k)$ of the Skorohod imbedding of shrunk random walks into the Wiener process as defined by* (65). *Then the sums below converge with probability* 1*:*

ITÔ INTEGRAL

(94)

$$\int_0^K f(W(s))\, dW(s) = \lim_{m\to\infty} \sum_{r=1}^{K2^{2m}} f(W(s_m(r-1)))\,(W(s_m(r)) - W(s_m(r-1))),$$

and

STRATONOVICH INTEGRAL

$$\int_0^K f(W(s)) \circ dW(s)$$

(95)

$$= \lim_{m\to\infty} \sum_{r=1}^{K2^{2m}} \frac{f(W(s_m(r-1))) + f(W(s_m(r)))}{2}\,(W(s_m(r)) - W(s_m(r-1))).$$

For the corresponding stochastic integrals we have the following formulae as well:

ITÔ FORMULA

$$\int_0^{W(K)} f(x)dx = \int_0^K f(W(s))dW(s) + \frac{1}{2}\int_0^K f'(W(s))ds, \tag{96}$$

and

STRATONOVICH FORMULA

$$\int_0^{W(K)} f(x)\; dx = \int_0^K f(W(s)) \circ dW(s). \tag{97}$$

PROOF. By the Itô case of Lemma 11 and the comments made after lemma, with probability 1, for all but finitely many m, we have the next equation for the sum in (94):

$$\sum_{r=1}^{\lfloor K/\Delta t\rfloor} \frac{f(W(s_m(r-1)))+f(W(s_m(r)))}{2}\,(W(s_m(r))-W(s_m(r-1)))$$

(98)

$$= T_{x=0}^{W(s_m(\lfloor K/\Delta t\rfloor))} f(x)\Delta x - \frac{1}{2}\sum_{r=1}^{\lfloor K/\Delta t\rfloor} \frac{f(W(s_m(r)))-f(W(s_m(r-1)))}{W(s_m(r))-W(s_m(r-1))}\Delta t,$$

where $\Delta x = 2^{-m}$ and $\Delta t = 2^{-2m}$.

For $t\in[0,K]$ set $t_m = \lfloor t/\Delta t\rfloor \Delta t$. Then $|t-t_m| \leqq \Delta t = 2^{-2m}$. By (93), $|t_m - s_m(\lfloor t/\Delta t\rfloor)| \leqq \sqrt{27Km}2^{-m}$ with probability 1 if m is large enough. This implies that

$$\max_{0\leqq t\leqq K} |t - s_m(\lfloor t/\Delta t\rfloor)| \to 0 \tag{99}$$

with probability 1 as $m\to\infty$. Further, the sample functions of the Wiener process being uniformly continuous on $[0,K']$ with probability 1, one gets that then

$$\max_{0\leqq t\leqq K} |W(t) - W(s_m(\lfloor t/\Delta t\rfloor))| \to 0 \tag{100}$$

as well.

Particularly, it follows that $W(s_m(\lfloor K/\Delta t\rfloor)) \to W(K)$ with probability 1 as $m\to\infty$. On the other hand, the trapezoidal sum $T_{x=0}^{a} f(x)\Delta x$ of a continuous function f is a Riemann sum corresponding to the partition $\{0, \frac{1}{2}\Delta x, \frac{3}{2}\Delta x, \dots, a-\frac{3}{2}\Delta x, a-\frac{1}{2}\Delta x, a\}$. Therefore the trapezoidal sums converge to $\int_{x=0}^{a} f(x)dx$ as $\Delta x\to 0$. These show that for any $\epsilon>0$,

$$\begin{aligned}
&\left|\int_0^{W(K)} f(x)dx - T_{x=0}^{W(s_m(\lfloor K/\Delta t\rfloor))} f(x)\Delta x\right| \\
&\leqq \left|\int_0^{W(s_m(\lfloor K/\Delta t\rfloor))} f(x)dx - T_{x=0}^{W(s_m(\lfloor K/\Delta t\rfloor))} f(x)\Delta x\right| \\
&\quad + \left|\int_{W(s_m(\lfloor K/\Delta t\rfloor))}^{W(K)} f(x)dx\right| \\
&< \epsilon/2 + \epsilon/2 = \epsilon
\end{aligned}$$

with probability 1 if m is large enough. That is, the trapezoidal sum in (98) tends to the corresponding integral with probability 1:

$$\lim_{m\to\infty} T_{x=0}^{W(s_m(\lfloor K/\Delta t\rfloor))} f(x)\Delta x = \int_0^{W(K)} f(x)dx. \tag{101}$$

Now let us turn to the second sum in (98). By the definition of the first passage times, $W(s_m(r)) - W(s_m(r-1)) = \pm 2^{-m} = \pm\Delta x$, which tends to 0 as $m \to \infty$. Hence

$$\begin{aligned}&\frac{f(W(s_m(r))) - f(W(s_m(r-1)))}{W(s_m(r)) - W(s_m(r-1))} = \\ &\frac{f(W(s_m(r)) \mp \Delta x) - f(W(s_m(r)))}{\mp\Delta x}.\end{aligned} \tag{102}$$

We want to show that this difference quotient gets arbitrarily close to $f'(W(r\Delta t))$ if m is large enough.

To this end, let us consider the following problem from calculus. If f is a continuously differentiable function, $x_m \to x$ and $\Delta x_m \to 0$ as $m \to \infty$, let us consider the difference of $f'(x)$ and $(f(x_m + \Delta x_m) - f(x_m))/\Delta x_m$. By the mean value theorem, the latter diference quotient is equal to $f'(u_m)$, where u_m lies between $x_m \to x$ and $x_m + \Delta x_m \to x$. Since f' is continuous, this implies that

$$\frac{f(x_m + \Delta x_m) - f(x_m)}{\Delta x_m} \to f'(x), \tag{103}$$

as $m \to \infty$.

In our present context, $x = W(t)$ and $x_m = W(s_m(\lfloor t/\Delta t\rfloor))$, where $0 \leq t \leq K$. Since the sample functions of $W(t)$ are continuous with probability 1, it follows from the max-min theorem that their ranges are contained in bounded intervals. Over such a bounded interval the function f' is uniformly continuous, therefore (99), (100), and (103) imply

(104)

$$\max_{0\leq t\leq K}\left| f'(W(t)) - \frac{f(W(s_m(\lfloor t/\Delta t\rfloor)) \mp \Delta x) - f(W(s_m(\lfloor t/\Delta t\rfloor)))}{\mp\Delta x}\right| \to 0$$

with probability 1 as $m \to \infty$. (Remember that now $\Delta x = 2^{-m}$ and $\Delta t = 2^{-2m}$.)

Particularly, for any $\epsilon > 0$, we have

$$\begin{aligned}&\max_{1\leq r\Delta t\leq K}\left|\frac{f(W(s_m(r))) - f(W(s_m(r-1)))}{W(s_m(r)) - W(s_m(r-1))} - f'(W(r\Delta t))\right| \\ &= \max_{1\leq r\Delta t\leq K}\left|\frac{f(W(s_m(r)) \mp \Delta x) - f(W(s_m(r)))}{\mp\Delta x} - f'(W(r\Delta t))\right| < \frac{\epsilon}{3K}\end{aligned} \tag{105}$$

with probability 1 assuming m is large enough.

The function $f'(W(s))$ is continuous with probability 1, so its Riemann sums over $[0, K]$ converge to the corresponding integral as the norm of the partition tends to 0. Thus by (105),

$$\begin{aligned}
&\left| \int_0^K f'(W(s))ds - \sum_{r=1}^{\lfloor K/\Delta t \rfloor} \frac{f(W(s_m(r))) - f(W(s_m(r-1)))}{W(s_m(r)) - W(s_m(r-1))} \Delta t \right| \\
&\leq \sum_{r=1}^{\lfloor K/\Delta t \rfloor} \left| f'(W(r\Delta t)) - \frac{f(W(s_m(r)) \mp \Delta x) - f(W(s_m(r)))}{\mp \Delta x} \right| \Delta t \\
&+ \left| \int_0^{K_m} f'(W(s))ds - \sum_{r=1}^{\lfloor K/\Delta t \rfloor} f'(W(r\Delta t))\Delta t \right| + \left| \int_{K_m}^{K} f'(W(s))ds \right| \\
&< \frac{\epsilon}{3K} K + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon,
\end{aligned}$$

with probability 1 if m is large enough. Here $K_m = \lfloor K/\Delta t \rfloor \Delta t$.

Therefore the second sum in (98) also tends to the corresponding integral with probability 1:

$$\lim_{m \to \infty} \frac{1}{2} \sum_{r=1}^{\lfloor K/\Delta t \rfloor} \frac{f(W(s_m(r))) - f(W(s_m(r-1)))}{W(s_m(r)) - W(s_m(r-1))} \Delta t = \frac{1}{2} \int_0^K f'(W(s))ds.$$

This proves that the defining sum of the Itô integral in (94) converges with probability 1 as $m \to \infty$, and for the limit we have Itô formula (96).

Also, by the Stratonovich case of Lemma 11 and the comments made after the lemma, with probability 1, for all but finitely many m, we have the following equation for the sum in (95):

$$\begin{aligned}
&\sum_{r=1}^{\lfloor K/\Delta t \rfloor} \frac{f(W(s_m(r-1))) + f(W(s_m(r)))}{2} (W(s_m(r)) - W(s_m(r-1))) \\
&\qquad = T_{x=0}^{W(s_m(\lfloor K/\Delta t \rfloor))} f(x)\Delta x.
\end{aligned}$$

We saw in (101) that this trapezoidal sum converges to the corresponding integral with probability 1 as $m \to \infty$. Therefore the defining sum of the Stratonovich integral in (95) converges as well, and for the limit we have formula (97). $\square$

Since the Itô and Stratonovich formulae are valid for the usual definitions of the corresponding stochastic integrals as well, this shows that the usual definitions agree with the definitions given in this paper.

As we mentioned in a special case, the interesting feature of Itô formula (96) is that it contains the non-classical term $\frac{1}{2}\int_0^K f'(W(s))ds$. If g denotes an antiderivative of the function f, then the Itô formula can be written as

$$g(W(t)) - g(W(0)) = \int_0^t g'(W(s))dW(s) + \frac{1}{2}\int_0^t g''(W(s))ds,$$

or formally as the following non-classical chain rule for differentials:

$$dg(W(t)) = g'(W(t))dW(t) + \frac{1}{2}g''(W(t))dt.$$

We mention that other, more complicated versions of Itô formula can be proved by essentially the same method, see [8]. Also, as shown there, multiple stochastic integrals can be defined analogously as the stochastic integrals defined above.

ACKNOWLEDGEMENTS. I am obliged to Pál Révész from whom I learned many of the ideas underlying this paper. Also, I am grateful to Gábor Tusnády for his valuable help and to Tamás Linder for several useful comments. Special thanks are due to the Informatics Department of the University of Pisa and particularly to Vincenzo Manca for their hospitality while writing the last part of the paper.

REFERENCES

[1] BREIMAN, L., *Probability*, Addison-Wesley, Reading, Mass., 1968. *MR* **37** #4841

[2] FELLER, W., *An introduction to probability theory and its applications*, Vol. I, Third ed., J. Wiley, New York, 1968. *MR* **37** #3604

[3] FELLER, W., *An introduction to probability theory and its applications*, Vol. II, J. Wiley, New York, 1966. *MR* **35** #1048. Second ed.: 1971. *MR* **42** #5292

[4] KNIGHT, F. B., *Essentials of Brownian motion and diffusion*, Mathematical Surveys, 18, American Mathematical Society, Providence, R.I., 1981. *MR* **82m**:60098

[5] PROTTER, P., *Stochastic integration and differential equations. A new approach*, Applications of Mathematics, 21, Springer, Berlin, 1990. *MR* **91i**:60148

[6] RÉVÉSZ, P., *Random walk in random and non-random environments*, World Scientific, Singapore, 1990. *MR* **92c**:60096

[7] RUDIN, W., *Principles of mathematical analysis*, Third ed., International Series in Pure and Applied Mathematics, McGraw-Hill, New York, 1976. *MR* **52** #5893

[8] SZABADOS, T., A discrete Itô's formula, *Limit theorems in probability and statistics* (Pécs, 1989), Colloq. Math. Soc. János Bolyai, **57**, North-Holland, Amsterdam, 1990, 491–502. *MR* **92i**:60105

(Received August 25, 1994)

BUDAPESTI MŰSZAKI EGYETEM
VILLAMOSMÉRNÖKI ÉS INFORMATIKAI KAR
MATEMATIKAI ÉS SZÁMITÁSTUDOMÁNYI TANSZÉK
EGRY JÓZSEF U. 20–22
H–1521 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 299–322

BOLTZMANN'S ERGODIC HYPOTHESIS, A CONJECTURE FOR CENTURIES?[1]

D. SZÁSZ[2]

Dedicated to P. Révész for his sixtieth birthday

Abstract

An overview of the history of Ludwig Boltzmann's more than one hundred year old ergodic hypothesis is given. The existing main results, the majority of which is connected with the theory of billiards, are surveyed, and some perspectives of the theory and interesting and realistic problems are also mentioned.

In 1964 Werner Heisenberg was elected a honorary doctor of Loránd Eötvös University, Budapest. In his inaugural lecture he made a point that sounded something like this: "A theoretical physicist feels best if there is no rigorously defined mathematical object behind his considerations". Certainly, Heisenberg was having the early years of quantum mechanics in his mind but what he said perfectly fitted the work of Ludwig Boltzmann as well. One could choose several areas of his interest to illustrate this statement, out of which the history of the ergodic hypothesis we are going to elaborate on is only one.[3]

As it was so nicely explained in Professor Gallavotti's illuminating lecture at this conference, G(1995), though the rigorously defined mathematical

1991 *Mathematics Subject Classification.* Primary 58F15; Secondary 58F11, 82B03.

Key words and phrases. Boltzmann's ergodic hypothesis, hard ball systems, Sinai's conjecture, billiards.

[1] Lecture presented on International Symposium in Honor of Boltzmann's 150th Birthday, February 23–26, 1994, Vienna.

[2] Research supported by the Hungarian National Foundation for Scientific Research Grant No. 1902.

[3] Another striking example, perhaps not sufficiently widely known, is the case with the Boltzmann equation. He published it in 1872, B(1872), and the first mathematically satisfactory derivation of the equation was only obtained more than 100 years later in 1975 by Oscar Lanford, L(1975), though the picture is still not complete. Thus needless to say that Boltzmann's original argument was highly intuitive. At the same time, however, it was so much challenging for the great mathematician, David Hilbert that he included among his celebrated collection of 23 problems presented at the International Mathematical Congress held at Paris in 1900 the sixth one with the title "Mathematical Treatment of the Axioms of Physics" (see H(1900)). In its formulation, besides requiring an axiomatic approach to the theory of probabilities, Hilbert also says: "it is therefore very desirable that the discussion of the foundations of mechanics be taken up by mathematicians also. Thus Boltzmann's work on the principles of mechanics suggests the problem of developing mathematically the limiting processes, there merely indicated, which lead from the atomistic view to laws of motion of continua". Boltzmann's law of motion of continua is, of course, his equation.

object behind Boltzmann's considerations around the ergodic hypothesis was indeed missing, Boltzmann was ingenious in inventing mathematical paradigmas and in mastering mathematical calculations on them to guess the truth and to obtain convincing power, and even without having the mathematical object he understood many things better than we do now.

1. Boltzmann's ergodic hypothesis

During the 1870s and 1880s, various forms of the ergodic hypothesis were used by Boltzmann in his works on the foundations of statistical mechanics (see e. g. B(1871) and B(1884); for a historic account also F(1989)). An advanced formulation of the hypothesis would sound as follows:

BOLTZMANN'S ERGODIC HYPOTHESIS. *For large systems of interacting particles in equilibrium time averages are close to the ensemble, or equilibrium average.*

REMARK. In this paper — with the exception of Section 10 — equilibrium averages always mean microcanonical ones, i. e. the Liouville measure on the submanifold of the phase space specified by the trivial invariants of the motion.

More precisely, if f is a measurement (i.e. a function on the phase space of the system), then as N, the size of the system (for instance, the number of particles) tends to infinity, then

$$\frac{1}{T}\int_0^T f(S^t x)dt \to \int f(x)d\mu(x) \tag{1}$$

where μ is the equilibrium measure, and $S^t x$ is the time evolution of the phase point x.

We immediately note that if N varies, then f and μ also depend on N, and thus, for a mathematically strict statement one ought to specify the sense of the convergence in (1), too. Let us look at the main steps of the history of Boltzmann's hypothesis — without intending to provide a complete account though I think such a study should be done. One major incompleteness of our survey is that it does not go into the history of the quasi-ergodic hypothesis at all; as to some recent results about it see H(1991) and Y(1992).

2. Finding a mathematical object, a notion and a problem (from Boltzmann to von Neumann, i.e. from 1870 until 1931)

It took quite a time until the mathematical object of the ergodic hypothesis was found. Indeed, only in 1929, Koopman, K(1931), began to investigate groups of measure-preserving transformations of a measure space or in other language, groups of unitary operators in a Hilbert space[4]. Koopman's idea was apparently in the air, and several mathematicians, including among others G. Birkhoff, M. S. Stone and A. Weil, contributed to the birth of ergodic theory; for a historic account see M(1990).

More precisely, let M be an abstract space, the phase space of the system and μ be a probability measure on (a σ-algebra of) M. The dynamics is a one-parameter group $S^{\mathbb{R}} = \{S^t : -\infty < t < \infty\}$ of measure preserving transformations, i.e. for every measurable subset $A \subset M$, and for every $t \in \in \mathbb{R}$ $\mu(S^{-t}A) = \mu(A)$.

Here, of course, μ is the equilibrium measure of the system. Let finally, $f : M \to \mathbb{R}$ be a measurement such that $f \in L_2(\mu)$. Thus the object [i. e. $(M, S^{\mathbb{R}}, d\mu)$ with the functions f] is defined.

In 1931, von Neumann proved the first ergodic theorem, the so called

Mean Ergodic Theorem (N(1932)). *As* $T \to \infty$,

$$\frac{1}{T}\int_0^T f(S^t x)dt \to \bar{f}(x)$$

in the L_2*-sense.*

(The exact story of the first ergodic theorems is explained in the note of Birkhoff and Koopman, B-K(1932).)[5]

The proof of the mean ergodic theorem is not difficult but it is worth noting that — even more than 20 years later — Neumann very highly appreciated exactly this achievement among his various findings in the vast territory of his interest. In 1954, when answering a questionnaire of the American Mathematical Society, his works on the ergodic theorems were named by himself among his most important discoveries (the other two were the mathematicial foundations of quantum mechanics, and further operator-algebras, called today Neumann-algebras).

The limiting function $\bar{f}(x)$ satisfies two further important properties:

(i) $$\bar{f}(x) = \mathsf{E}(f/\mathcal{I}),$$

[4] This progress was preceded by the success of Lebesgue's theory of measure which, on another path, also led, in 1933, Kolmogorov to the laying down the axiomatic foundations of probability theory.

[5] Though his name is not explicitly mentioned, Boltzmann's influence on von Neumann is also seen in the title of his earlier work on quantum ergodic theory, N(1929): "Beweis des Ergodensatzes und des H-Theorems in der neuen Mechanik".

where $\mathcal{I}$ is the σ-algebra of the invariant sets, or in other words $\bar{f}$ is the projection of f onto the subspace of functions invariant with respect to the dynamics $S^{\mathbb{R}}$;

$$\text{(ii)} \qquad \int_M f d\mu = \int_M \bar{f} \mu \quad \text{whenever} \quad f \in L_2.$$

An extremely important consequence is the following: if the only invariant functions are the constants or, in other words, there are only trivial invariant sets (i.e. $\mu((S^{-t}A \setminus A) \cup (A \setminus S^{-t}A)) = 0$ implies $\mu(A) = 0$ or 1), then, first of all, $\bar{f}$ is a constant for every f and, moreover, by (ii), $\bar{f} = \int f d\mu$. Consequently, the ergodic theorem says that, then as $T \to \infty$,

$$\frac{1}{T} \int_0^T f(S^t x) dt \to \int_M f d\mu \tag{2}$$

in the L_2-sense.

This statement is much reminiscent to Boltzmann's hypothesis but here we still have just one fixed system and not various ones for different values of N. Anyway, define the system to be *ergodic*, if the only invariant functions are the constants. Then we know that, for ergodic systems, the relation (2), i.e. a version of the ergodic hypothesis holds.

Summarizing: we have a mathematical model (groups of measure preserving transformations), the notion of ergodicity and, finally, the problem of establishing the ergodicity of a system we are interested in from the mechanical point of view.

We note that, a bit later in 1931, Birkhoff, B(1931) (and also Khintchine) could, moreover, prove that the convergence in (2) holds almost everywhere as well.

This progress led to the birth of an independent branch of mathematics: ergodic theory. This theory then began his autonomous evolution within mathematics and several sub-branches were also born. Just to mention some, one of them studies various forms and generalizations of the ergodic theorems, another one stronger forms of stochasticity, a special branch — quite interesting for our present discussion — investigates the ergodicity of particular systems, among them those arising from mechanics, a further one the isomorphism problem of various dynamical systems,etc.

3. Proving the first relevant theorem (from Neumann to Sinai, from 1931 until 1970)

The methods for establishing the ergodicity of mechanical systems came from a different though related domain, from the theory of dynamical systems. In 1938–39, Hedlund, He(1939) and Hopf, Ho(1939) found a method for demonstrating the ergodicity of geodesic flows on compact manifolds of negative curvature. Their main conceptual discovery was that the so called hyperbolic behaviour of dynamical systems could imply and, in fact, did imply ergodicity in the aforementioned models.

Hyperbolicity means, in other words, instability, i.e. the exponential divergence of trajectories starting arbitrarily close to each other, or else sensitivity to the initial conditions. The simplest example of a hyperbolic system is Arnold's famous cat, the linear automorphism of the torus (cat stands for a Continuous Automorphism of the Torus) . Indeed, if we consider the map T_A of the 2-torus $\mathbb{R}^2 \mid \mathbb{Z}^2$ onto itself defined by the (hyperbolic) matrix $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, then we see that the image of the cat gets expanded in one direction and contracted in a transversal one, with the expansion (contraction) being the strongest in the eigendirection of the matrix corresponding to the eigenvalue $\lambda_u > 1 (\lambda_s < 1)$.

In 1942, very soon after Hedlund's and Hopf's fundamental results, the Russian physicist, N. S. Krylov discovered that systems of elastic hard balls show an instability similar to the one observed at geodesic flows on manifolds with negative curvature, cf. K(1942). This finding and the progress of the ideas of Hedlund and Hopf in the theory of hyperbolic dynamical systems justified Sinai's stronger version of Boltzmann's ergodic hypothesis formulated in 1963 for the particular system of elastic hard balls.

The Boltzmann-Sinai Ergodic Hypothesis (S(1963)). *The system of N hard balls given on $\mathbb{T}^2$ or $\mathbb{T}^3$ is ergodic for any $N \geqq 2$.*

Since mechanical systems also have conserved quantities, this conjecture is understood so that ergodicity is expected to hold on (connected components of) the submanifold of the phase space specified by the invariants of motion.

The conceptual surprise of this conjecture compared to Boltzmann's original formulation was that no large N was assumed. In fact, ergodicity (and further stronger mixing properties, like the K-property) was expected to hold for any fixed $N \geqq 2$!

In 1970, Sinai, S(1970) was able to verify this conjecture in the case of $N = 2$ 2-dimensional discs moving on the 2-torus $\mathbb{T}^2$.

Before giving an insight into Sinai's approach, let us mention the limitations of this nice ergodic behaviour for systems with a fixed number of degrees of freedom.

4. Appearance of non-ergodic behaviour (consequences of the KAM-theory, 1954–1974)

In nature, we have important examples of systems of interacting particles (or bodies) that are stable and not unstable like systems of hard balls.

The most striking example is — fortunately — the solar system. The fact that it consists of bodies of different masses is not of great importance, more significant is the fact that here the interaction is different.

The year 1954 brought two important discoveries. Kolmogorov's 1954 work, K(1954) and its later evolution — thanks first of all to the achievements of Arnold and Moser (in particular, A(1963) and M(1962)) in the 60's — indicated that we may well have a situation when invariant tori with dimension half of that of the phase space can fill a set of positive measure (we note that in completely integrable systems such invariant tori do foliate the whole phase space). Another, not so explicit, warning came from the numerical work of Fermi-Pasta-Ulam, F-P-U(1955) demonstrating that the asymptotic equipartition of the energy of modes may fail. As to a detailed exposition of this experiment and its effects we refer to the survey H(1983).

In the 1974 work of Markus–Meyer, M-M(1974) summarizing the previous progress there were two important statements out of which the first one is more remarkable for our discussion.

Theorem. *In the space of smooth Hamiltonians*

(1) *The nonergodic ones form a dense open subset;*

(2) *The nonintegrable ones contain a countable intersection of dense open subsets (i.e. form a subset of second Baire category).*

Without going into technical details we note that the statements are formulated in the C^∞-topology and a Hamiltonian is called ergodic if, for almost every values of the energy, the system is ergodic on the corresponding submanifold of the phase space.

Thus, for generic Hamiltonians, we cannot expect ergodicity, and in the final sections of the paper we will return to the question of what kind of ergodic behaviour can then be expected for them. The forthcoming discussion will be focused on the comparatively simple case of hard ball systems.

5. Sinai's setup. Billiards (1970)

We start with a simple trick traditional both in mathematics and physics: instead of treating N particles we consider just one particle in a high dimensional phase space. More concretely: Let us assume, in general, that a system of $N(\geqq 2)$ balls of unit mass and radii $r > 0$ are given on $\mathbb{T}^\nu$, the ν-dimensional unit torus ($\nu \geqq 2$). Denote the phase point of the i'th ball by $(q_i, v_i) \in \mathbb{T}^\nu \times \mathbb{R}^\nu$. The configuration space $\mathbf{Q}$ of the N balls is a subset of $\mathbb{T}^{N\nu}$: from $\mathbb{T}^{N\nu}$ we cut out $\binom{N}{2}$ cylindric scatterers:

$$\tilde{C}_{i,j} = \left\{ Q = (q_1, \dots, q_N) \in \mathbb{T}^{N\nu} : \mid q_i - q_j \mid < 2r \right\},$$

$1 \leqq i < j \leqq N$. The energy $H = \frac{1}{2} \sum_1^N v_i^2$ and the total momentum $P = \sum_1^N v_i$ are first integrals of the motion. Thus, without loss of generality, we can assume that $H = \frac{1}{2}$ and $P = 0$ and, moreover, that the sum of spatial components $B = \sum_1^N q_i = 0$ (if $P \neq 0$, then the center of mass has an additional conditionally periodic or periodic motion). For these values of H, P and B, the phase space of the system reduces to $M := \mathbf{Q} \times \mathcal{S}_{N\nu-\nu-1}$ where

$$\mathbf{Q} := \left\{ Q \in \tilde{\mathbf{Q}} \setminus \bigcup_{1 \leqq i < j \leqq N} \tilde{C}_{i,j} : \sum_1^N q_i = 0 \right\}$$

with $d := \dim \mathbf{Q} = N\nu - \nu$, and where $\mathcal{S}_k$ denotes, in general, the k-dimensional unit sphere. It is easy to see that the dynamics of the N balls, determined by their uniform motion with elastic collisions on one hand, and the billiard flow $\{S^t : t \in \mathbb{R}\}$ on $\mathbf{Q}$ with specular reflections on $\partial\mathbf{Q}$ on the other hand, are isomorphic and they conserve the Liouville measure $d\mu = \text{const} \cdot dq \cdot dv$.

We recall that a *billiard* is a dynamical system describing the motion of a point particle in a connected, compact domain $\mathbf{Q} \subset \mathbb{R}^d$ or $\mathbf{Q} \subset \mathbb{T}^d = \text{Tor}^d$, $d \geqq 2$ with a piecewise C^2-smooth boundary. Inside $\mathbf{Q}$ the motion is uniform while the reflection at the boundary $\partial\mathbf{Q}$ is elastic (the angle of reflection equals the angle of incidence, cf. Figure 1). Since the absolute value of the velocity is a first integral of motion, the phase space of our system can be identified with the unit tangent bundle over $\mathbf{Q}$. Namely, the configuration space is $\mathbf{Q}$ while the phase space is $M = \mathbf{Q} \times S_{d-1}$ where S_{d-1} is the surface of the unit d-ball. In other words, every phase point x is of the form (q, v) where $q \in \mathbf{Q}$ and $v \in S_{d-1}$. The natural projections $\pi : M \to \mathbf{Q}$ and $p : M \to S_{d-1}$ are defined by $\pi(q, v) = q$ and by $p(q, v) = v$, respectively.

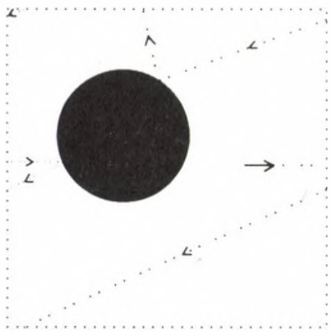

Fig. 1

Suppose that $\partial\mathbf{Q} = \cup_1^k \partial\mathbf{Q}_i$ where $\partial\mathbf{Q}_i$ are the smooth components of the boundary. Denote $\partial M = \partial\mathbf{Q} \times S_{d-1}$ and let $n(q)$ be the unit normal vector of the boundary component $\partial\mathbf{Q}_i$ at $q \in \partial\mathbf{Q}_i$ directed inwards $\mathbf{Q}$. In billiards, isomorphic to hard ball systems, the scatterers are convex cylinders if $N \geqq 3$, and are (strictly convex) balls if $N = 2$. The observation of Krylov and Sinai was that a billiard with strictly convex scatterers behaves like a hyperbolic dynamical system, whereas in one with just convex scatterers there is some partial hyperbolicity. We will illustrate this observation after some definitions.

We say that a billiard is *dispersing* (a Sinai-billiard) if each $\partial \mathbf{Q}_i$ is strictly convex, and we say it is *semi-dispersing* if each $\partial \mathbf{Q}_i$ is convex. The billiards in Figures 2 and 3 are dispersing. Indeed, they correspond to the system of two discs on $\mathbb{T}^2$; the first one to the case $R < 1/4$ and the second one to the case $1/4 < R < 1/2$.

The third one is a semi-dispersing billiard given on $\mathbb{T}^3$ with two cylindric scatterers. This paradigm was the first semi-dispersing but not dispersing billiard whose ergodicity was established (cf. K-S-Sz(1989)).

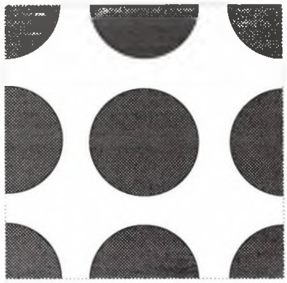

Fig. 2

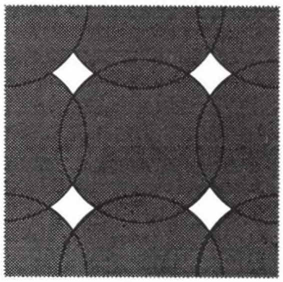

Fig. 3

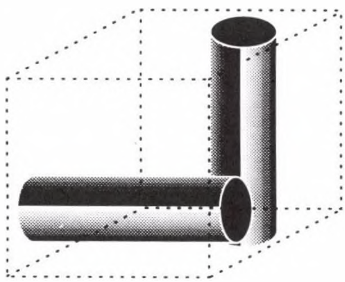

Fig. 4

The mechanism producing hyperbolicity in a dispersing billiard can be seen the best on Figure 5 borrowing the illustration from optics. Assume we have a strictly convex scatterer on $\mathbb{T}^d$ and imagine it is a mirror. Take, $x = (Q, V) \in M$, and the codimension one hyperplane Γ through Q in the configuration space perpendicular to the velocity V. By attaching to points of Γ velocities identical to V we obtain a wavefront $\tilde{\Gamma}$ in the phase space M. After one reflection from the mirror scatterer, our wavefront gets strictly convex while the linear distances measured on Γ get uniformly expanded. This mechanism is exactly the one providing the (uniform) hyperbolicity of a dispersing billiard.

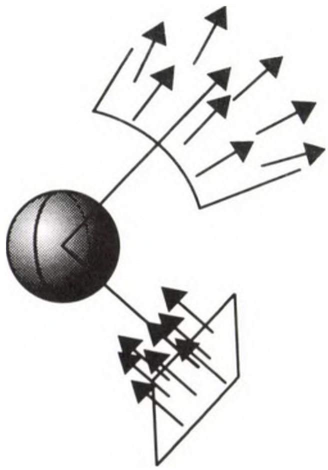

Fig. 5

Sinai's 1970 work used the theory of uniformly hyperbolic smooth dynamical systems which had had an intensive progress in the 60s and culminated in the 1967 paper of Anosov and Sinai, A-S(1967). The serious difficulty Sinai had to cope with was that billiards were not smooth dynamical systems. Indeed, if a smooth wavefront gets reflected from a scatterer and it contains a tangency, then though the reflected wavefront will be continuous, its second derivative will have a jump at the tangency. This circumstance causes serious technical difficulties: in smooth uniformly hyperbolic dynamical systems the stable and unstable invariant manifolds, the fundamental tools of the theory are smooth and unbounded, whereas in billiards their smooth components can be arbitrarily small.

6. $N = 2$ balls (1970–1987). Local ergodicity of semi-dispersing billiards

As mentioned earlier, Sinai, in 1970, in his celebrated paper obtained the first rigorous result in relation to the Boltzmann-Sinai ergodic hypothesis: he could show that $N = 2$ discs on the 2-torus $\mathbb{T}^2$ was a K-system.

In fact, his result was formulated for $2 - D$ dispersing billiards (Sinai-billiards) with a finite horizon. A billiard has *finite horizon* if there is no collision-free trajectory in it. This condition is fulfilled by a two-billiard if $R > \frac{1}{4}$ (cf. Figure 3). In this case the configuration space consists of four connected components, and, of course, ergodicity is claimed on each of them. For the case of $R < \frac{1}{4}$ (cf. Figure 2), a $2-D$ billiard with infinite horizon, the corresponding result was proved by Bunimovich and Sinai in 1973, B-S(1973). On the basis of their work it was understood that a $2-D$ dispersing billiard was ergodic.

A multidimensional generalization of their theorem was only obtained in 1987. Indeed, Chernov and Sinai, S-Ch(1987) were, in general, investigating semi-dispersing billiards and introduced the basic notion of *sufficiency* of an orbit or equivalently of a phase point. The main consequence of sufficiency is that, in a suitably small neighbourhood of a sufficient point, the system is hyperbolic, though not uniformly. Next we present this notion in its minimal form as suggested in K-S-Sz (1990).

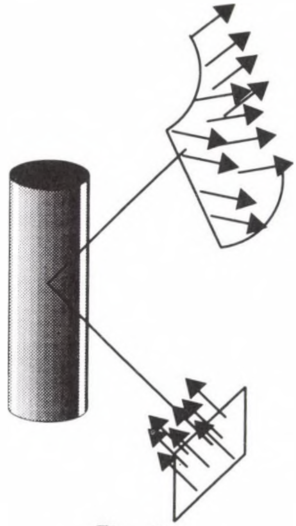

Fig. 6

Our starting point is Figure 6, similar to Figure 5. It shows that, if a scatterer is not strictly convex but just convex, like e. g. a cylinder, then the image of the hyperplanar wavefront Γ with parallel velocities will not be curved in the directions parallel with the constituent subspace of the cylinder, but in the transversal directions, only. However, the uncurved neutral directions can still die out after several reflections on differently oriented cylindric (or, in general, convex) scatterers.

Now for the definition of sufficiency. Assume that $S^{[a,b]}x$ is a finite trajectory segment, which is regular, i.e. it avoids singularities.

Let $S^a x = (Q, V) \in M$ and consider the hyperplanar wavefront $\bar{\Gamma}(S^a x) :=$ $= \{(Q + dQ, V) : dQ \text{ small } \in \mathbb{R}^d \text{ and } \langle dQ, V\rangle = 0\}$ (by denoting $\pi(x) = Q$ for $x = (Q, V)$ we see that, indeed, $\pi(\bar{\gamma})$ is part of a hyperplane).

We say that the trajectory segment $S^{[a,b]}x$ is *sufficient* if $\pi(S^b\bar{\Gamma})$ is strictly convex (see Figure 7). (To obtain a geometric or optical feeling of this notion, the reader is again suggested to imagine mirror-surfaced scatterers.)

A phase point $x \in M$ is sufficient if its trajectory is sufficient (i.e. it contains a sufficient trajectory segment). In physical terms, sufficiency of a trajectory segment means that, during the time interval $[a, b]$, the trajectory of x encounters all degrees of freedom of the system.

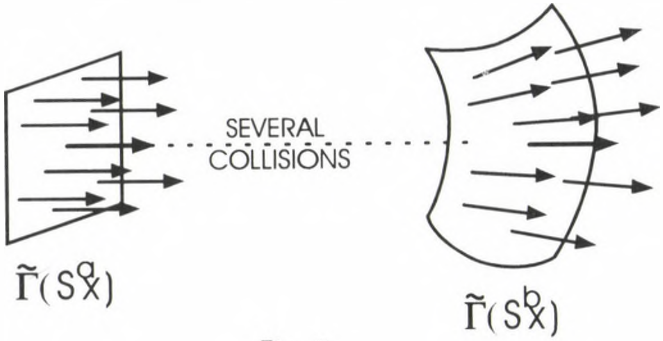

Fig. 7

If a trajectory segment is not sufficient, then the curvature of $\pi(S^b\tilde{\Gamma})$ at $\pi(S^b x)$ necessarily vanishes in certain directions forming the so-called *neutral subspace*. Simple geometric considerations (cf. K-S-Sz(1990)) show that a sufficient trajectory segment generates an expansion rate uniformly larger than 1 in some neighbourhood of the point $S^a x$. Then, by Poincaré recurrence and the ergodic theorem, it is not hard to see that, in some neighbourhood of $S^a x$, the relevant Lyapunov exponents of the system are not zero. In other words, in this neighbourhood, the system is hyperbolic. This observation should motivate the non-trivial

FUNDAMENTAL THEOREM FOR SEMI-DISPERSING BILLIARDS (S-Ch-(1987)). *Assume that a semi-dispersing billiard satisfies some geometric conditions and the Chernov — Sinai Ansatz, a condition strongly connected with the singularities of the system.*

If $x \in M$ is a sufficient point, then it has an open neighbourhood U in the phase space belonging to one ergonent (i.e. ergodic component).

(A simplified and suitably generalized version of this theorem, the so-called 'transversal fundamental theorem' was given in K-S-Sz(1990). Moreover, a version of the fundamental theorem formulated for symplectic maps with singularities can be found in L-W(1994).) The property expressed in the statement is usually called *local ergodicity*. If almost every phase point of a semi-dispersing billiard is sufficient, then, of course, it may have at most a countable number of ergonents. In some cases it is not hard then to derive the *global ergodicity* of the system, i.e. to show that there is just one ergonent in the phase space. Note that it also follows from the general theory that, on each ergonent, the system is Kolmogorov mixing. A much important consequence is thus the following

Corollary (S-Ch(1987)). *Every dispersing billiard is ergodic, and, moreover, is a K-flow. In particular, the system of $N = 2$ balls on the ν-torus is a K-flow if $r < \frac{1}{2}$.*

(For details, see K-S-Sz(1990).)

7. $N \geqq 3$ balls (1989–). Global ergodicity of semi-dispersing billiards

With the fundamental theorem for semi-dispersing billiards in mind, the proof of their global ergodicity boils down to

(1) first demonstrating the Chernov–Sinai Ansatz, an important condition of the fundamental theorem, and
(2) to then showing that the subset of non-sufficient points is a topologically small set of measure zero; for instance, its topological codimension is not smaller than two.

In Sz(1993), we gave a sketch of the strategy worked out in our papers with A. Krámli and N. Simányi for the core part, and here we will just list the main results obtained so far.[6]

(1) in 1991, Krámli, Simányi and the present author, K-S-Sz(1991) demonstrated the K-property of $N = 3$ balls on the ν–torus whenever $\nu \geqq 2$;
(2) in 1992, again the previous authors, K-S-Sz(1992) improved their methods to get the ergodicity of $N = 4$ balls on the ν–torus whenever $\nu \geqq 3$;
(3) in 1992, Simányi, S(1992) was able to establish the so far strongest result for hard ball systems: the system of $N \geqq 2$ balls is ergodic on the ν–torus whenever $\nu \geqq N$; his method is based on his Connecting Path Formula characterizing the neutral subspace of a trajectory segment.

The configuration of the cylindric scatterers of a billiard isomorphic to a hard ball system inherits the permutation symmetry of the balls. A natural generalization of hard ball systems is to investigate cylindric billiards in general, i. e. billiards with solely cylinders as scatterers. To this end consider compact affine subspaces $L^i : 1 \leqq i \leqq N$, $N \geqq 1$ in the d-torus $\mathbf{T}^d$ (with $\dim\, L^i \leqq d - 2$), and denote $C^i := \{Q := (q_1, \ldots, q_d) : \mathrm{dist}(Q, L^i) \leqq r^i\}$, $1 \leqq$ $\leqq i \leqq N$ where each r^i is positive. The billiard in $\mathbf{Q} := \mathbf{T}^d \setminus (\cup_{i=1}^N C^i)$ is *a billiard with cylindric scatterers.*

For cylindric billiards the following results have been obtained:

(1) in 1989, Krámli, Simányi and the present author, K-S-Sz(1989) considered a 3-dimensional orthogonal cylindric billiard (cf. Figure 4); they obtained its K-property and thus this was the first semi-dispersing — but not dispersing — billiard whose ergodicity was shown.

[6] Most recently, in the Summer of 1994, Simányi and Szász were able to prove Sinai's hypothesis for hard ball systems with a restricted graph of interactions.

(2) in 1993, motivated by a question of John Mather, the present author started a systematic study of cylindric billiards and found a sufficient and necessary condition for the ergodicity of a class of them: for *orthogonal cylindric billiards*, cf. Sz(1993), Sz(1994). These are characterized by the property that the constituent subspace of any cylindric scatterer is spanned by some of the coordinate vectors adapted to the orthogonal coordinate system where $\mathbf{T}^d$ is given;

(3) in 1994, Simányi and the present author, S-Sz(1994) found necessary and sufficient conditions for the K-property of a toric billiard with two arbitrary cylindric scatterers.

Since the class of cylindric billiards is relatively simple, one can hope for general necessary and sufficient conditions for the ergodicity (and the K-property) of these systems. Indeed, we next formulate a conjecture containing a general sufficient condition.

CONJECTURE (Szász, 1992). *Assume that the configuration domain of a cylindric billiard is connected, and no pairs of the scatterers are tangent. If there is at least one sufficient point, then the billiard is K.*

8. The Boltzmann–Sinai ergodic hypothesis in pencase type models

In order to resolve some difficulties on the way to establishing the Boltzmann–Sinai ergodic hypothesis, Chernov and Sinai, S-Ch(1985) suggested the study of a quasi-one-dimensional model of hard balls. It is given on an elongated torus of the type $(L\mathrm{T}^1) \times \mathrm{T}^{\nu-1}$ where L is a sufficiently large number compared to R (see Figure 8). The main assumption is

$$\frac{\sqrt{\nu-1}}{4} < R < \frac{1}{2}$$

ensuring that the order of balls (in the direction of $L\mathrm{T}^1$) is invariant under the dynamics. Thus the model, which was called by Chernov and Sinai a *pencase*, is realizable if $2 \leqq \nu \leqq 4$. If we number the balls in their order: $1, 2, \dots, N$, then a particular feature of the model is that only the pairs of consecutive balls (i.e. $\{1,2\}, \{2,3\}, \dots, \{N,1\}$) can interact.

Fig. 8

The first result for a pencase type model was reached in 1992 by Bunimovich–Liverani–Pellegrinotti–Sukhov. Instead of a torus their model lives in a domain with dispersing boundaries (see Figure 9) and the sizes of the domain ensure that

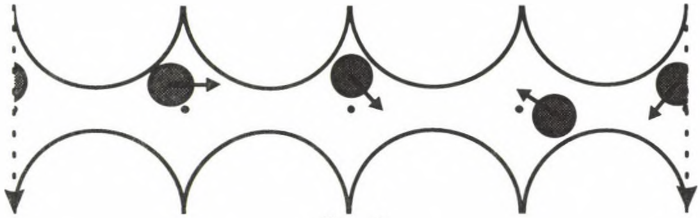

Fig. 9

(1) each ball is restricted to a fundamental domain of the "pencase" (the throats between them are smaller than $2R$);
(2) between consecutive collisions of a particular ball, it should always hit a dispersing boundary;
(3) the pairs of balls in neighbouring domains can, indeed, interact.

A billiard of this type is realizable in arbitrary dimension and the results of the aforementioned authors was that the system was K. This particular model was, in fact, the first one where the Boltzmann–Sinai ergodic hypothesis got settled for any N and $\nu \geqq 2$.

THEOREM (B-L-P-S(1992)). *The B-L-P-S pencase is a K-system for any* $N, \nu \geqq 2$.

For some time it seemed so that the proof of ergodicity for the original Chernov–Sinai pencase was not easier than that for general hard ball systems. Nevertheless, — with Nándor Simányi — we could recently demonstrate the following

THEOREM (S-Sz(1994-B)). *The Chernov–Sinai pencase is a K-system for any* $N \geqq 2, \nu = 4$. *If* $\nu = 3$, *the system has open ergodic components.*

The restriction $\nu \neq 2$ seems, at present, important whereas that of $\nu < 5$ only arises since the model, as invented by its authors, does not exist for $\nu \geqq 5$. One could, however, introduce less realistic models that do exist for $\nu \geqq 5$, too, and for them our proof would also work but we do not want to stay on them.

There is, however, another, more natural way to introduce models with a pencase-type interaction in any dimension. Consider, namely, N balls, numbered $1, 2, \ldots, N$ on the unit torus $\mathbb{T}^\nu$. The restriction is that only pairs of balls with neighbouring numbers, i.e. again only the pairs $\{1,2\}, \{2,3\}, \ldots, \{N-1, N\}, \{N, 1\}$ interact while other pairs can go through each other. This billiard is, of course, again a cylindric one.

THEOREM (S-Sz(1994-B)). *The system with pencase-type interaction is a K-system whenever* $N \geqq 2, \nu \geqq 4$. *If* $\nu = 3$, *the system has open ergodic components.*

(Froeschlé(1978) (cf. H(1983)) introduced the notion of *connectivity* as the ratio of the number of particles a given particle can interact with and of the number of all particles. His experiments suggested that this ratio can be related to the good ergodic properties of a system; in particular, below a critical value of the connectivity, a significant fraction of the phase space is occupied with invariant tori. Our theorem shows, however, that, for hard ball systems, the ergodic behaviour already appears at a connectivity arbitrarily close to zero.)

9. Ergodicity of systems with a fixed number of degrees of freedom

From the work of Markus–Meyer mentioned in Section 4 we know that ergodic Hamiltonians are in a sense exceptional. Nevertheless, it makes sense to look for possibly more of them since the mechanisms occurring in these can also help to understand the onset of chaotic behaviour, for instance, the appearance of a large ergodic component in nonergodic systems.

In Sections 5-7 we discussed billiard systems. Here we mention three classes of Hamiltonians, for which Donnay and Liverani, D-L(1991) could, in 1991, demonstrate ergodicity. These are systems of $N=2$ particles on $\mathbf{T}^2$ interacting via a rotation-invariant pair potential $V(r)$. These system have the same conserved quantities as the system of two hard discs and we assume that $v_1^2+v_2^2=1$, $v_1+v_2=0$, $q_1+q_2=0$. We do not give here all the conditions since we are mainly interested in the qualitative description of these interactions.

Assume in all cases that for some $R>0$

(1) $V(r)=0$ if $r \geq R$;

(2) $V(r) \in C^2(O,R)$;

(3) $\lim_{r \to 0} r^2 V(r)=0$;

(4) for $h(r)=r^2(1-2V(r))$, and for except one value of $r \in (O,R)$ $h'(r)>0$.

Potentials in the first class are repelling ones (see Figure 10). The additional condition besides (1)–(4) is now

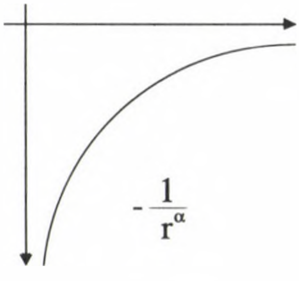

Fig. 10

(5) $V(R-)=0$ and $V'(R-)<0$.

Then, under some more conditions, the system is K. As it is evident from the conditions, V is, though continuous, not C^1 at $r=R$ (see Figure 11). Indeed, the jump of V' in R as required by (5) plays the same role as the effect of a reflection in a dispersing billiard. This phenomenon was first observed by Kubo in 1976 (K(1976)), and he, and later he and Murata, K-M(1981) could already establish the K- and the B-property of such systems under more restrictive conditions than those of Donnay and Liverani. It is a natural question whether the Kubo-type singularity can also lead to ergodicity in the case of several particles. In fact, we recall the following

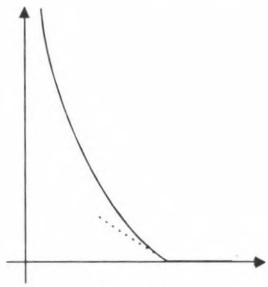

Fig. 11

PROBLEM (Liverani–Szász, 1990). *Let $N=3$, $\nu=2$. Is it possible to find a Kubo-type interaction (i.e. one satisfying the conditions* (1)–(5) *formulated before) such that the system is ergodic?*

A simpler problem could be the generalization of the Kubo–Donnay–Liverani result for the case $N=2$, $\nu \geqq 3$ though, as observed by Wojtkowski, W(1990-C), in the multidimensional case new, unpleasant phenomena may arise.

The second class investigated by Donnay and Liverani contains attracting potentials. In 1987, Knauf, K(1987) showed that for attracting interactions with singularities at $r=0$ of the type $-\frac{1}{r^{2(1-\frac{1}{n})}}$, $n=2,3,4,\ldots$, the system was ergodic. Donnay and Liverani's main achievement was that they could get rid of the assumption that n was an integer (see Figure 10). Their main condition besides (1)–(4) is

(6) $V'(r) \geqq 0$ if $\nu \in (0,R)$ and $V(R)=V'(R)=0$.

From the conceptual point of view the most remarkable is their third class since here the potential is everywhere smooth. The basic feature of interactions in the third class is that, for some $r_c<R$, the circle of radius r_c is a closed orbit (see Figure 12). Interestingly enough this orbit plays the role of a singularity.

In all cases, the existence of potentials satisfying the aforelisted conditions is proved. For a given potential satisfying the appropriate requirements then ergodicity is fulfilled at sufficiently high energy. It is worth noting that having proved first that the Lyapunov exponent is non-zero, the proof of ergodicity can be obtained by a suitable adaptation of the fundamental theorem for semi-dispersing billiards (cf. Section 6).

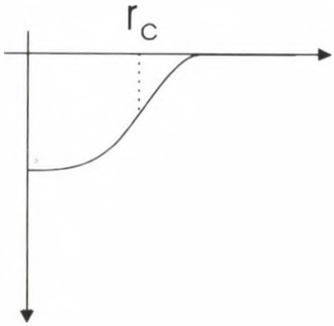

Fig. 12

An interesting class of models was introduced and studied by Wojtkowski, W(1990-A) and W(1990-B). Here a one-dimensional system of N particles of different masses moves in an external field, and the interaction is elastic collision. The non-vanishing of Lyapunov exponents has been proved in several cases, but establishing global ergodicity still seems to be difficult.

10. Ergodicity of systems with an increasing number of degrees of freedom

The situation when the number of particles increases exactly corresponds to Boltzmann's original question which — in modern terminology — could sound as follows: find, for a generic Hamiltonian, the asymptotic behaviour in the thermodynamic limit. This question is still not formulated precisely. From the various possible ways, the right one should, of course, be selected as dictated by the main applications. At present, as it seems to me, a very important application should be in the field of the derivation of hydrodynamic equation from microscopic, Hamiltonian principles. It is clear that, the so far strongest method worked out in the last decade by Varadhan and his coworkers, O-V-Y(1993) would require a form related to Boltzmann's hypothesis but we can still not select the right form (we note that the results obtained until now are valid for stochastic systems and not for purely Hamiltonian ones).

The conceptually simplest and widemost known form of a hypothesis is the following: denote as before the number of particles by N, and by $p(N)$ the relative measure of the phase space occupied by invariant tori. For simplicity, the interaction is fixed and $\frac{V}{N} = \text{const}$ (for definiteness, we assume that the system lives on the torus $V^{\frac{1}{\nu}}\mathbf{T}^{\nu}$). Then the first conjecture is that $p(N) \to 0$ as $N \to \infty$. A stronger conjecture would then require that the complement to the set of invariant tori contains a large ergodic component whose measure gets close to one.

Hénon (1983) and Galgani (1985) discussed in detail the situation and the connection of these conjectures to the one on the limiting equipartition of energy between the modes of the system . The conclusion is that the picture is not clear at all. There are, on one hand, interactions when numerical work of Froeschlé and Scheidecker (1975) indicates that $p(N) \to 0$ as expected. They investigated a one-dimensional model with the Hamiltonian

$$H = \frac{1}{2\sigma}\sum_1^N p_i^2 + 2\pi G\sigma^2 \sum_{i<j} | q_i - q_j | .$$

On the other hand, the famous Fermi–Pasta–Ulam (1955) experiment supported doubts about the conjectures by detecting the failure of the limiting equipartition of energy. This was also a one-dimensional model with the Hamiltonian

$$H = \frac{1}{2}\sum_1^N p_i^2 + \sum_1^{N-1} V(q_{i+1} - q_i),$$

where $V(q) = \frac{1}{2}q^2 + \alpha q^3$.

These works generated a vivid interest in the problem. For the contradictory views about it, the reader is suggested to consult the aforementioned papers of Hénon and Galgani and for a more recent review that of Galgani–Giorgili–Martinolli–Vanzini (1993).

As I learnt from Gregory Eyink, for establishing hydrodynamic equations in the sense of the approach of Varadhan's method, a weaker form of the conjecture would also be sufficient. It is not necessary to have one large ergodic component. It seems that a weakly increasing upper bound for the number of ergodic components and, of course, a good upper bound on $p(N)$ could be sufficient. The picture here, however, needs more elaboration and the problems seem very difficult.

11. Ergodicity of systems with an infinite number of degrees of freedom

Since the situation with large but finite systems is so complicated, I expect that the solution of equilibrium statistical physics should be borrowed. Whereas even a rigorous definition of a phase transition in a finite system — not speaking about its demonstration — is not an easy task, the question gets much simpler for infinite systems. In my view, first the ergodicity of infinite systems should be understood.

The very first result for an infinite system was obtained in 1971 by Sinai and Volkovysky for the ideal gas: it was shown to be a K-system (V-S(1971)). (A weaker result was obtained by Dobrushin already in 1956, see D(1956).) For the first glance this sounds as a surprise since in the ideal gas there is no velocity mixing at all. Indeed, in the formulation of ergodicity one should be a bit cautious. By denoting the phase space by $M = \{\{(q_i, v_i) : i \in \mathbb{Z}\} : \{q_i\}$ is locally finite $\}$, the equilibrium measure is $P_\lambda(\{q_i : i \in \mathbb{Z}\}) \otimes \prod F(dv_i)$ where P_λ is a Poisson measure with density λ and $F(dv)$ is an arbitrary non-degenerate probability distribution in $\mathbb{R}^\nu$, and ergodicity holds with respect to this invariant measure. The proof reveals an apparently new mechanism of ergodicity : mixing — understood, of course, in time — is the result of the initial spatial mixing. In other words: the equilibrium measure is Poisson, i.e. a measure with independent increments. Now as time proceeds, in a fixed box of our observation, particles starting from more and more distant intervals appear and their numbers are, roughly speaking, independent. This phenomenon can be proved to provide mixing in time.

The same observation was used, in more delicate arguments, for showing the K-property for different variants of the Rayleigh-gas, among others, by Goldstein–Lebowitz–Ravishankar (1982), Boldrighini–Pellegrinotti–Presutti–Sinai–Soloveychik (1984) and L. Erdős–Tuyen (1991). In these models only one particle interacts with all the other ones and the equilibrium measure is still Poisson. A related model is the Lorentz gas where — similarly to the ideal one — there is no interaction between the particles, but the dynamics of each particle obeys a strong mixing in space. Based upon this mixing, Sinai demonstrated the K-property in S(1979).

Now a problem which I find very interesting and quite realistic is the following one:

PROBLEM (Szász, 1990). *Consider an infinite pencase obtained as $N \to \infty$ of the finite ones was introduced in Section 8. Prove that the natural Gibbs measure is ergodic. (Here, of course, the possible values of the dimension are $\nu = 2, 3, 4$.)*

(Infinite models also raise the question of existence of the dynamics, but for this model it was answered affirmatively by Alexander, A(1976).) In the proof of ergodicity two mechanisms can be exploited: the hyperbolic behaviour of the interaction as done for the finite pencase or the spatial mixing of the equilibrium distribution as in cases of the ideal gas or the Rayleigh one. At present, however, I do not see an easy way for any of these possibilities, in particular, for the second one. For the first one it is a natural idea to start building up the hyperbolic theory of infinite-dimensional dynamical systems and trying to define, for instance, the stable and unstable invariant manifolds first, and then to prove their existence.

In Section 10 we already mentioned the problem of the derivation of hydrodynamic type equations. A sufficient condition in some cases for the method of O-V-Y(1993) to work is the following ergodicity type condition: *every "regular" state, invariant with respect to both translations in the space and the dynamics, is a mixture of canonical Gibbs measures.* This property is apparently stronger than ergodicity, but, as remarked in F-F-L(1994), to prove such an ergodicity for deterministic Hamiltonian systems is still a formidable unsolved problem. (In fact, Fritz, Funaki and Lebowitz verify this property for a random Hamiltonian system, and their paper is, moreover, also recommended for further reference.) We add that regularity above means that the state has finite relative entropy (per unit volume) with respect to the Gibbs measure. This assumption implies, in particular, that the conditional distribution in any finite volume Λ, given the configuration outside Λ, is absolutely continuous with respect to the Lebesgue measure.

12. Concluding remarks

In this lecture I have been concentrating on the history of Boltzmann's ergodic hypothesis. I think that the second half of the title of the talk is already justified if we focus our interest to just the question of ergodicity. After more than one hundred years, ergodicity is still not established in the simplest mechanical model, in the system of elastic hard balls though I expect we are not far from a solution. But as to generic interactions, even the questions are not clearly posed and it might well be that there will not be a final understanding after the next hundred years either. And we have not touched upon more delicate, physically fundamental properties for whose proofs one should refine the methods used in studying ergodicity of the system involved. Without aiming at completeness we just mention the problems

(1) of the decay of correlations (cf. Ch(1994); here and in the forthcoming cases only the last reference, I am aware of, will be provided, where further ones can also be found),
(2) of the convergence to equilibrium, K-Sz(1983),
(3) of the calculation of and bounds on the entropy of mechanical systems, Ch(1991),
(4) and, finally, of the recurrence properties of such systems, K-Sz(1985).

Acknowledgements. Sincere thanks are due to G. Eyink and J. Fritz for their suggestions about the relation of ergodicity and the hydrodynamic limit transition. I am also much grateful to M. Herman and B. Weiss for their helpful remarks on the history of the ergodic hypothesis and some references.

References

[A(1976)] Alexander, R., Time evolution for infinitely many hard spheres, *Comm. Math. Phys.* **49** (1976), 217–232. *MR* **55** #7252

[A-S(1967)] Anosov, D. V. and Sinai, Ya. G., Certain smooth ergodic systems, *Uspehi Mat. Nauk* **22** (1967), no. 5 (137), 107–172 (in Russian). *MR* **37** #370

[A(1963)] Arnold, V. I., Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian, *Uspehi. Mat. Nauk* **18** (1963), no. 5 (113), 13–40 (in Russian). *MR* **29** #328

[Bi(1931)] Birkhoff, G. D., Proof of the ergodic theorem, *Proc. Nat. Acad. Sci. U.S.A.* **17** (1931), 656–660. *Zbl* **3**, 256

[Bi-K(1932)] BIRKHOFF, G. D. and KOOPMAN, B. O., Recent contributions to ergodic theory, *Proc. Nat. Acad. Sci. U.S.A.* **18** (1932), 279–282. *Zbl* **4**, 311

[B-P-P-S-S(1985)] BOLDRIGHINI, C., PELLEGRINOTTI, A., PRESUTTI, E., SINAI, YA. G. and SOLOVEICHIK, M. R., Ergodic properties of a semi-infinite one-dimensional system of statistical mechanics, *Comm. Math. Phys.* **101** (1985), 363–382. *MR* **87f**:82002

[B(1871)] BOLTZMANN, L., Einige allgemeine Sätze über das Wärmegleichgewicht, *Wien. Ber.* **63** (1871), 679–711. *Jahrbuch Fortschritte Math.* **3**, 536

[B(1872)] BOLTZMANN, L., Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen, *Wien. Ber.* **66** (1872), 275–370. *Jahrbuch Fortschritte Math.* **4**, 566–567

[B(1884)] BOLTZMANN, L., Über die Eigenschaften monocyklischer und anderer damit verwandter Systeme, *Kronecker J. = J. Reine Angew. Math.* **98** (1884), 68–94. *Jahrbuch Fortschritte Math.* **16**, 863

[B-L-P-S(1992)] BUNIMOVICH, L. A., LIVERANI, C., PELLEGRINOTTI, A. and SUKHOV, YU. M., Ergodic systems of n balls in a billiard table, *Comm. Math. Phys.* **146** (1992), 357–396. *MR* **93e**:58107

[B-S(1973)] BUNIMOVICH, L. A. and SINAI, YA. G., The fundamental theorem of the theory of scattering billiards, *Mat. Sb. (N. S.)* **90(132)** (1973), 415–431 (in Russian). *MR* **51** #3395

[CH(1991)] CHERNOV, N. I., A new proof of Sinai's formula for entropy of hyperbolic billiards. Its application to Lorentz gas and stadium, Joint Inst. Nuclear Res., Dubna, 1990. *MR* **91k**:58103; A new proof of the Sinai's formula for computing the entropy of hyperbolic billiards. Application to Lorentz gas and to the Bunimovich stadium, *Funktsional Anal. i Prilozhen.* **25** (1991), no. 3, 50–69 (in Russian). Translation in *Functional Anal. Appl.* **25** (1991), 204–219. *MR* **93c**:58110

[CH(1994)] CHERNOV, N. I., Limit theorems and Markov approximations for chaotic dynamical systems, 1994, 1–45 (manuscript).

[D(1956)] DOBRUSHIN, R. L., On Poisson's law for distribution of particles in space, *Ukrain. Mat. Ž.* **8** (1956), 127–134 (in Russian). *MR* **18**-341

[D-L(1991)] DONNAY, V. and LIVERANI C., Potentials on the two-torus for which the Hamiltonian flow is ergodic, *Comm. Math. Phys.* **135** (1991), 267–302. *MR* **92a**:58082

[E-T(199?)] ERDÖS, L. and TUYEN, D. Q., Ergodic properties of the multidimensional Rayleigh gas with a semipermeable barrier, *J. Stat. Phys.* (to appear).

[F-P-U(1955)] FERMI, E., PASTA, J. and ULAM, S., *Studies of nonlinear problems*, Los Alamos Report LA–1940, 1955.

[F(1989)] FLAMM, D., Boltzmann's statistical approach to irreversibility, UWThPh **1989-4** (1989), 1–11.

[F-F-L(1994)] FRITZ, J., FUNAKI, T. and LEBOWITZ, J. L., Stationary states of random Hamiltonian systems, *Probab. Theory Related Fields* **99** (1994), 211-236.

[F(1978)] FROESCHLÉ, C., *Phys. Rev. A* (3) **18** (1978), 277.

[F-S(1975)] FROESCHLÉ, C. and SCHEIDECKER, J.-P., *Phys. Rev. A* (3) **12** (1975), 2137.

[G(1985)] GALGANI, L., Ordered and chaotic motions in Hamiltonian systems and the problem of energy partition, *Chaos in astrophysics*, ed. by J. R. Buchler, 1985, 245–257.

[G-G-M-V(1993)] GALGANI, L., GIORGILLI, A., MARTINOLI, A. and VANZINI, S., On the problem of energy equipartition for large systems of the Fermi–Pasta–Ulam type: analytical and numerical estimates, *Phys. D* **59** (1992), 334–348. *MR* **93h**:58135

[G(1995)] GALLAVOTTI, G., Ergodicity, ensembles, irreversibility in Boltzmann and beyond, *J. Stat. Phys.* **78** (1995), 1571–1589.

[G-L-R(1982)] GOLDSTEIN, S., LEBOWITZ, J. L. and RAVISHANKAR, K., Ergodic properties of a system in contact with a heat bath: a one-dimensional model, *Comm. Math. Phys.* **85** (1982), 419–427. *MR* **84k**:82053

[HE(1939)] HEDLUND, G. A., The dynamics of geodesic flows, *Bull. Amer. Math. Soc.* **45** (1939), 241–260. *Zbl* **20**, 403–404

[H(1983)] HÉNON, M., Numerical exploration of Hamiltonian systems, *Chaotic behavior of deterministic systems* (Les Houches, 1981), North-Holland, Amsterdam–New York, 1983, 53–170. *MR* **84k**:58074

[H(1991)] HERMAN, M. R., Stabilité topologique des systèmes dynamiques conservatifs, 1991, pp. 15 (manuscript).

[H(1900)] HILBERT, D., Mathematische Probleme, *Gött. Nachr.* 1900, 253–297. *Jahrbuch Fortschritte Math.* **31**, 68–69

[Ho(1939)] HOPF, E., Statistik der geodätischen Linien in Mannigfaltigkeiten negativer Krümmung, *Ber. Verh. sächs. Akad. Leipzig* **91** (1939), 261–304. *Zbl* **24**, 80–81

[K(1987)] KNAUF, A., Ergodic and topological properties of Coulombic periodic potentials, *Comm. Math. Phys.* **110** (1987), 89–112. *MR* **88f**:58119

[K(1954)] KOLMOGOROV, A. N., On conservation of conditionally periodic motions for a small change in Hamilton's function, *Dokl. Akad. Nauk SSSR (N. S.)* **98** (1954), 527–530 (in Russian). *MR* **16**-924

[K(1931)] KOOPMAN, B. O., Hamiltonian systems and transformations in Hilbert space, *Proc. Nat. Acad. Sci. U.S.A.* **17** (1931), 315–318. *Zbl* **2**, 57

[K-S-Sz(1989)] KRÁMLI, A., SIMÁNYI, N. and SZÁSZ, D., Ergodic properties of semi-dispersing billiards I. Two cylindric scatterers in the 3D torus, *Nonlinearity* **2** (1989), 311–326. *MR* **90f**:58138

[K-S-Sz(1990)] KRÁMLI, A., SIMÁNYI, N. and SZÁSZ, D., A "transversal" fundamental theorem for semi-dispersing billiards, *Comm. Math. Phys.* **129** (1990), 535–560. *MR* **91b**:58125; see also *MR* **92e**:58119

[K-S-Sz(1991)] KRÁMLI, A., SIMÁNYI, N. and SZÁSZ, D., The K-property of three billiard balls, *Ann. of Math.* (2) **133** (1991), 37–72. *MR* **92b**:58130

[K-S-Sz(1992)] KRÁMLI, A., SIMÁNYI, N. and SZÁSZ, D., The K-property of four billiard balls, *Comm. Math. Phys.* **144** (1992), 107–148. *MR* **93a**:58107

[K-Sz(1983)] KRÁMLI, A. and SZÁSZ, D., On the convergence to equilibrium of the Lorentz gas, *Functions, series, operators*, Vol. I–II (Budapest, 1980), Colloq. Math. Soc. J. Bolyai, Vol. 35, North-Holland, Amsterdam–New York, 1983, 757–766. *MR* **85i**:82011

[K-Sz(1985)] KRÁMLI, A. and SZÁSZ, D., The problem of recurrence for Lorentz processes, *Comm. Math. Phys.* **98** (1985), 539–552. *MR* **86m**:60063

[K(1942)] KRYLOV, N. S., The processes of relaxation of statistical systems and the criterion of mechanical instability, Thesis, 1942. In: *Development of Krylov's ideas*, Princeton University Press, Princeton, NJ, 1977, 193–238.

[K(1976)] KUBO, I., Perturbed billiard systems. I. The ergodicity of the motion of a particle in a compound central field, *Nagoya Math. J.* **61** (1976), 1–57. *MR* **55**#6486

[K-M(1981)] KUBO, I. and MURATA, H., Perturbed billiard systems. II. Bernoulli properties, *Nagoya Math. J.* **81** (1981), 1–25. *MR* **83b**:58043

[L(1975)] LANFORD, O. E., Time evolution of large classical systems, *Dynamical systems, theory and applications* (Recontres, Battelle Res. Inst., Seattle, Wash., 1974), Lecture Notes in Phys., Vol. 38, Springer, Berlin, 1975, 1–111. *MR* **57** #18653

[L-W(1994)] LIVERANI, C. and WOJTKOWSKI, M., Ergodicity in Hamiltonian systems, *Dynamics Reported* (to appear).

[M(1990)] MACKEY, G. W., Von Neumann and the early days of ergodic theory, *The legacy of John von Neumann* (Hempstead, NY, 1988), Proc. Sympos. Pure Math., 50, American Mathematical Society, Providence, RI, 1990, 25–38. *MR* **92a**:01056. See also: *The scope and history of commutative and noncommutative harmonic analysis*, History of Mathematics, 5, American Mathematical Society, Providence, RI; London Mathematical Society, London, 1992, 311–324. *MR* **93g**:22006

[M-M(1974)] MARKUS, L. and MEYER, K. R., Generic Hamiltonian dynamical systems are neither integrable nor ergodic, *Memoirs Amer. Math. Soc.*, No. 144, American Math. Society, Providence, RI, 1974. *MR* **56** #9588

[M(1962)] MOSER, J., On invariant curves of area-preserving mappings of an annulus, *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl.* II 1962, 1–20. *MR* **26** #5255

[N(1929)] NEUMANN, J. v., Beweis des Ergodensatzes und des H-Theorems in der neuen Mechanik, *Z. f. Physik* **57** (1929), 30–70. *Jahrbuch Fortschritte Math.* **55**, 523

[N(1932)] NEUMANN, J. v., Proof of the quasi-ergodic hypothesis, *Proc. Nat. Acad. Sci. U.S.A.* **18** (1932), 70–82. *Zbl* **4**,310–311

[O-V-Y(1993)] OLLA, S., VARADHAN, S. R. S. and YAU, H. T., Hydrodynamic limit for a Hamiltonian system with weak noise, *Comm. Math. Phys.* **155** (1993), 523–560.

[S(1992)] SIMÁNYI, N., The K-property of N billiard balls. I, *Invent. Math.* **108** (1992), 521–548. *MR* **93m**:58066a; The K-property of N billiard balls. II. Computation of neutral linear spaces, *Invent. Math.* **110** (1992), 151–172. *MR* **93m**:58066b

[S-Sz(1994-A)] SIMÁNYI, N. and SZÁSZ, D., The K-property of 4D billiards with non-orthogonal cylindric scatterers, *J. Statist. Phys.* (to appear).

[S-Sz(1994-B)] SIMÁNYI, N. and SZÁSZ, D. (in preparation).

[S(1984)] SIMON, B., Fifteen problems in mathematical physics, *Perspectives in mathematics*, Birkhäuser, Basel–Boston, Mass., 1984, 423–454. *MR* **86g**:81003

[S(1963)] SINAI, YA. G., On the foundations of the ergodic hypothesis for a dynamical system of statistical mechanics, *Dokl. Akad. Nauk SSSR* **153** (1963), 1261–1264 (in Russian). Translated as *Soviet Math. Dokl.* **4** (1963), 1818–1822. *RŽ Mat* 1964 #5 B497. *MR* **35** #5576

[S(1970)] SINAI, YA. G., Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards, *Uspehi Mat. Nauk* **25** (1970), no. 2 (152), 141–192 (in Russian).

[S-CH(1985)] SINAI, YA. G. and CHERNOV, N. I., Ergodic properties of some systems of 2-D discs and 3-D spheres, 1985 (manuscript).

[S(1979)] SINAI, YA. G., Ergodic properties of a Lorentz gas, *Funktsional. Anal. i Prilozhen.* **13** (1979), no. 3, 46–59 (in Russian). *MR* **81b**:28018

[S-CH(1987)] SINAI, YA. G. and CHERNOV, N. I., Ergodic properties of some systems of two-dimensional disks and three-dimensional balls, *Uspekhi Mat. Nauk* **42** (1987), 153–174. *MR* **89c**:58097

[Sz(1993)] SZÁSZ, D., Ergodicity of classical hard balls, *Physica A* **194** (1993), 86–92.

[Sz(1994)] SZÁSZ, D., The K-property of "orthogonal" cylindric billiards, *Comm. Math. Phys.* **160** (1994), 581–597.

[V-S(1971)] VOLKOVISSKII, K. L. and SINAI, YA. G., Ergodic properties of an ideal gas with an infinite number of degrees of freedom, *Funkcional. Anal. i Prilozen.* **5** (1971), no. 3, 19–21 (in Russian). *MR* **44** #6289

[W(1990-A)] WOJTKOWSKI, M. P., A system of one-dimensional balls with gravity, *Comm. Math. Phys.* **126** (1990), 507–533. *MR* **91h**:58066

[W(1990-B)] WOJTKOWSKI, M. P., The system of one-dimensional balls in an external field. II, *Comm. Math. Phys.* **127** (1990), 425–432. *MR* **91k**:58101

[W(1990-C)] WOJTKOWSKI, M. P., Linearly stable orbits in 3-dimensional billiards, *Comm. Math. Phys.* **129** (1990), 319–327. *MR* **91b**:58128

[Y(1992)] YOCCOZ, J-CH., Travaux de Herman sur les tores invariants, Séminaire Bourbaki, Vol. 1991/92, 44; *Astérisque* No. 206 (1992), Exp. No. 754, 311–344. *MR* **94g**:58195

(Received September 1, 1994)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 323–353

WEIGHTED APPROXIMATIONS OF PARTIAL SUM PROCESSES IN $D[0,\infty)$. I

B. SZYSZKOWICZ

Dedicated to Pál Révész for his sixtieth birthday

Abstract

Let $X_1, X_2, \ldots$ be independent, identically distributed random variables with $S(k) = X_1 + \cdots + X_k$. We prove best possible weighted approximations of $n^{-1/2}S(nt)$, $0 \leqq t \leqq 1$, by a standard Wiener process $\{W(t),\ 0 \leqq t < \infty\}$ assuming the existence of two moments only for X_1 and, consequently, obtain the weighted version of Donsker's theorem in $D[0,1]$ for the optimal class of weight functions. Considering functions h on $[1,\infty)$ such that $\limsup_{t\to\infty} |W(t)|/h(t) < \infty$ a.s. enables us to prove weighted approximations, and hence also weak convergence, of weighted partial sum processes in $D[1,\infty)$. In this case the admissible class of weight functions will be seen to be bigger than that for asymptotics on $[0,1]$. We show also that the class of weight functions for weighted sup- and L_p-functionals to converge in distribution is larger than that for weak convergence. Our proofs are based on Theorem of Major [17].

1. Introduction

Let $X_1, X_2, \ldots$ be independent, identically distributed random variables (i.i.d.r.v.'s) with $\mathbf{E}X_1 = 0$, $\mathbf{E}X_1^2 = 1$, partial sums $S(n) = X_1 + \cdots + X_n$ and let $\{W(t),\ 0 \leqq t < \infty\}$ denote the standard Wiener process starting at zero. After Donsker's theorem, the question arises under what conditions does weak convergence continue to hold for weighted partial sum processes $n^{-1/2}S(nt)/q(t)$, $0 < t \leqq 1$, where $q(t)$ is a nonnegative function on $(0,1]$ approaching zero as $t \to 0$. Assuming the existence of two moments only, in Section 2 we prove weighted approximations of partial sum processes $n^{-1/2}S(nt)$ by a standard Wiener process for the optimal class of weight functions, which is the same as the class of functions for which

$$\lim_{t\downarrow 0} |W(t)|/q(t) = 0 \ \text{ a.s..} \tag{1.1}$$

The motivation for studying weighted partial sums comes from earlier studies of Rényi [20], Chibisov [1], Pyke and Shorack [19], O'Reilly [18]

1991 *Mathematics Subject Classification.* Primary 60F17, 60F25; Secondary 60G50.
Key words and phrases. Partial sums, Wiener processes, the $D[0,\infty)$ weighted approximations, supremum metrics, L_p-distance.

Research supported by an NSERC Canada Grant at Carleton University, Ottawa.

and others concerning the asymptotic behaviour of weighted empirical and quantile processes. O'Reilly [18] proved the weak convergence of weighted partial sum processes in $C(0,1]$ under the assumption of $\mathbf{E}|X_1|^3 < \infty$.

Based on a refinement of the Csörgő and Révész [8] inequality for the strong approximation of the uniform empirical and quantile processes by a sequence of Brownian bridges, which in turn is based on the Komlós, Major and Tusnády KMT [14] inequality for the strong approximation of partial sum processes by a Wiener process, Csörgő, Csörgő, Horváth and Mason [2] established approximations of empirical and quantile processes by sequences of Brownian bridges in weighted supremum metrics for the optimal class of weight functions. A short proof of the basic approximation result was given by Csörgő and Horváth [3]. Using their method of proof together with the KMT [14] and Major [15] strong approximations of the partial sum process, one can construct a standard Wiener process $\{W(t),\ 0 \leqq t < \infty\}$ such that (cf. Csörgő and Horváth [4])

$$\sup_{\lambda/n \leqq t \leqq 1} n^{\nu-1/2}|S(nt) - W(nt)|/t^{1/2-\nu} = O(1) \quad \text{a.s.} \tag{1.2}$$

for any $0 \leqq \nu \leqq 1/2 - 1/r$ and $0 < \lambda < \infty$, if we assume

$$\mathbf{E}|X_1|^r < \infty \quad \text{for some} \quad r > 2.$$

While this gives weighted approximations for the optimal class of weight functions as in [2], r in (1.2) cannot be taken to be equal to 2. Indeed, a weighted embedding like that of (1.2) with $\nu = 0$ is impossible for $r = 2$. It would, for example, contradict Major [16]. Consequently, a new method of proof had to be developed for obtaining weighted approximations of $n^{-1/2}S(nt)$ under the assumption of two moments only. The proof of our Theorem 2.1 is based on a strong approximation theorem of Major [17], approximating the partial sum process of i.i.d.r.v.'s having only two moments by a Gaussian but not a Wiener process. As a corollary to our Theorem 2.1, we obtain the optimal weighted version of Donsker's theorem in supremum metrics. These results were announced in Szyszkowicz [23] and summarized in detail in Szyszkowicz [24], [26]. It is interesting to note here that the class of admissible weight functions q for the convergence in distribution of the sup-functional of $n^{-1/2}S(nt)/q(t)$ is bigger than for the weak convergence of the whole process in supremum norm. Such a phenomenon was first noticed and proved for weighted empirical and quantile processes by [2], and then for partial sums when assuming the existence of more than two moments by Csörgő and Horváth [4]. Here we prove such a result under the assumption of two moments only. In the light of these advances it is only natural to expect that the class of weight functions for the convergence in distribution of the L_p-functionals will be even bigger. For optimal weighted L_p-approximations of partial sum processes in $D[0,1]$ we refer to Szyszkowicz [25], [26], [28].

Obviously, all the results for $t \in [0,1]$ can be stated on $[0,T]$ for any $0 < T < \infty$. Since $W(t) \to \infty$ as $t \to \infty$, there is no weak convergence of $n^{-1/2}S(nt)$ on $[1,\infty)$. However, introducing appropriate weight functions opens up the possibility of studying such phenomena near infinity. In Section 3 we study asymptotics of weighted partial sum processes $n^{-1/2}S(nt)/h(t)$, $1 \leq t < \infty$, where $h(t)$ is a non-negative function on $[1,\infty)$ and $h(t) \to \infty$ as $t \to \infty$. While our weighted L_p-approximations on $[1,\infty)$ will be complete analogs of those on $[0,1]$, the optimal class of weight functions for weak convergence in $D[1,\infty)$ in the sup-norm will be seen to be bigger than the corresponding class for the weak convergence in $D[0,1]$. This is due to the fact that on $[1,\infty)$ there is no *a priori* need for a Chibisov–O'Reilly type theorem, that is to be guaranteed by $\lim_{t\to\infty} |W(t)|/h(t) = 0$ a.s., an analog of (1.1) near infinity. Namely, we obtain approximations in probability, and hence also weak convergence of our weighted partial sum processes in $D[1,\infty)$ whenever

$$\limsup_{t\to\infty} |W(t)|/h(t) < \infty \quad \text{a.s.}. \tag{1.3}$$

Considering q on $(0,1]$ (resp. h on $[1,\infty)$) which are positive and such that $q(t)$ is non-decreasing near zero (resp. $h(t)/t$ is non-increasing near infinity), we use the analytic conditions for (1.1) and (1.3) in terms of appropriate integrals as in [2] and Csörgő, Shao and Szyszkowicz [10], where the additional conditions of continuity of q (resp. of h) and monotonicity of $q(t)/t^{1/2}$ near zero (resp. of $h(t)/t^{1/2}$ near infinity) inherited from the classical Kolmogorov test for upper and lower class functions of a Wiener process are dropped. In the case of assuming the existence of more than two moments, an analog of (1.2) results in a complete solution of the problem in $D[1,\infty)$ (cf. Theorem 3.1 and Corollary 3.1). Such a method of proof, however, cannot work in case of having only two moments due to the fact that, similarly as on $[0,1]$, an analog of (1.2) is again impossible. Nevertheless, we succeed in proving approximations in probability and weak convergence in this case as well with weight functions satisfying (1.3), provided that $h(t)/t^{1/2}$ is slowly varying at infinity. Under the weaker assumption of monotonicity near infinity of $h(t)$ only (cf. the two sentences right after Corollary 3.2), when having two moments, we arrive at a Chibisov–O'Reilly type theorem near infinity (cf. Theorem 3.3 and Corollary 3.3).

In addition to considering the supremum norm, we obtain also optimal results for L_p-approximations, $0 < p < \infty$, on $[1,\infty)$. Again, the class of admissible weight functions will be seen to be bigger than in the case of sup-norm approximation.

While the results on $[0,1]$ in Section 2 are improvements of known results, all the results on $[1,\infty)$ of Section 3 are believed to be new. They were announced in Szyszkowicz [27].

For a treatize on recent advances on weighted approximations in probability and statistics in general, we refer to Csörgő and Horváth [6]. For the convergence of weighted partial sum processes in Banach function spaces we refer to the forthcoming paper by Csörgő, Norvaiša and Szyszkowicz. Asymptotic results for weighted partial sum processes are important in statistical applications, for example, in studying change-point problems (cf., e.g, Csörgő and Horváth [4]; Szyszkowicz [26]). For further results along these lines we refer to the forthcoming book on change-point analysis by Csörgő and Horváth.

2. Weighted approximations on $[0,1]$

Let Q be the class of functions q defined on $(0,1]$ which are positive, i.e.,

$$\inf_{\delta \leqq t \leqq 1} q(t) > 0 \qquad \text{for } 0 < \delta < 1 \tag{2.1}$$

and non-decreasing in a neighbourhood of zero. Using terminology introduced in Csörgő, Shao and Szyszkowicz [10], such a function q will be called a local function of a standard Wiener process $\{W(t),\ 0 \leqq t < \infty\}$ if

$$\limsup_{t \downarrow 0} |W(t)|/q(t) < \infty \quad \text{a.s.}. \tag{2.2}$$

A local function q of a standard Wiener process W will be called a Chibisov–O'Reilly local function of W if

$$\lim_{t \downarrow 0} |W(t)|/q(t) = 0 \quad \text{a.s.}. \tag{2.3}$$

Introduce the following integrals:

$$E(q,c) := \int_0^1 t^{-3/2} q(t) \exp(-ct^{-1}q^2(t))dt,$$

and

$$I(q,c) := \int_0^1 t^{-1} \exp(-ct^{-1}q^2(t))dt,$$

for some constant $0 < c < \infty$.

The integral $E(q,c)$ appeared in the works of Kolmogorov, Petrovski, Erdős and Feller. For details we refer to Itô and McKean ([12], Section 1.8).

The integral $I(q,c)$ appeared in the works of Chibisov [1] and O'Reilly [18].

For further comments on these two integrals, as well as for the proof of the next three theorems, we refer to [2], (cf. also Csörgő, Shao and Szyszkowicz [10]). We have (cf. Proposition 3.1, and Theorems 3.3 and 3.4, respectively, of [2]):

THEOREM A. (i) *Whenever the integral* $I(q,c)<\infty$ *for* $q\in Q$, *then* $E(q,c+\epsilon)<\infty$ *for every* $\epsilon>0$ *and* $q(t)/t^{1/2}\to\infty$ *as* $t\downarrow 0$.

(ii) *Whenever* $E(q,c)<\infty$ *and* $q(t)/t^{1/2}\to\infty$ *as* $t\downarrow 0$ *for* $q\in Q$, *then* $I(q,c)<\infty$.

THEOREM B. *A function* $q\in Q$ *is a local function of a standard Wiener process starting at zero if and only if the integral* $I(q,c)<\infty$ *for some* $c>0$ *or, equivalenty, if and only if the integral* $E(q,c)<\infty$ *for some* $c>0$ *and* $\lim_{t\downarrow 0} q(t)/t^{1/2}=\infty$.

THEOREM C. *A function* $q\in Q$ *is a Chibisov–O'Reilly local function of a standard Wiener process if and only if the integral* $I(q,c)<\infty$ *for all* $c>0$ *or, equivalently, if and only if the integral* $E(q,c)<\infty$ *for all* $c>0$ *and* $\lim_{t\downarrow 0} q(t)/t^{1/2}=\infty$.

By Lemma 4.4.4 of Csörgő and Révész [9] (cf. also Section A.2 in Csörgő and Horváth [6]), we can assume without loss of generality that our probability space $(\Omega, A, \mathbf{P})$ accommodates all random variables and stochastic processes introduced so far and later on.

THEOREM 2.1. *Let* $X_1, X_2, \ldots$ *be independent, identically distributed random variables such that*

$$\mathbf{E}X_1=0, \quad \mathbf{E}X_1^2=1, \tag{2.4}$$

and for each $n\geqq 1$ *let* $S(nt)=\sum_{i=1}^{[nt]} X_i$. *Then a standard Wiener process* $\{W(t),\ 0\leqq t<\infty\}$ *can be constructed in such a way that the following statements hold true.*

(a) *Let* $q\in Q$. *Then, as* $n\to\infty$

$$\sup_{0<t\leqq 1} \left| n^{-1/2}(S(nt)-W(nt))\right| \Big/ q(t)=o_P(1) \tag{2.5}$$

if and only if $I(q,c)<\infty$ *for all* $c>0$.

(b) *Let* $q\in Q$. *Then, as* $n\to\infty$

$$\sup_{0<t\leqq 1} \left| n^{-1/2}(S(nt)-W(nt))\right| \Big/ q(t)=O_P(1) \tag{2.6}$$

if and only if $I(q,c)<\infty$ *for some* $c>0$.

REMARK 2.1. According to Theorem C, the statement (2.5) of Theorem 2.1 is also equivalent to

$$E(q,c)<\infty \text{ for all } c>0, \text{ and } \lim_{t\downarrow 0} q(t)/t^{1/2}=\infty.$$

Similarly, according to Theorem B, the statement (2.6) of Theorem 2.1 is also equivalent to

$$E(q,c)<\infty \text{ for some } c>0, \text{ and } \lim_{t\downarrow 0} q(t)/t^{1/2}=\infty.$$

COROLLARY 2.1. *Let $X_1, X_2, \ldots$ be i.i.d.r.v.'s such that*

$$\mathbf{E}X_1=0, \quad \mathbf{E}X_1^2=1.$$

Then, as $n\to\infty$, with $q\in Q$ we have

$$n^{-1/2}S(nt)/q(t)\xrightarrow{\mathcal{D}}W(t)/q(t) \text{ in } D[0,1]$$

if and only if $I(q,c)<\infty$ for all $c>0$, where W is a standard Wiener process.

This weak convergence follows from (a) of Theorem 2.1 on account of $S(nt)=0, 0\leq t<1/n$, and by Theorem C. Throughout this paper weak convergence statements on Skorohod spaces are stated as corollaries to approximations in probability. Naturally, when talking about weighted weak convergence on such spaces we will always assume that the weights are c.d.l.g. functions.

Part (a) of Corollary 2.1 was proved by O'Reilly [18] for continuous $q\in Q$ under the assumption of $\mathbf{E}|X_1|^3<\infty$. He asserted also that the third moment condition could be dropped so that his theorem would remain true, but no proof was given. The proof given below for Theorem 2.1 appears to be a first, and certainly a new one, under the assumption of two moments only for X_1. We will use the following result of Major [17].

THEOREM D. *Let a distribution $F(x)$ be given with $\int x dF(x)=0$, $\int x^2 dF(x)=1$. Define*

$$\sigma_k^2=\int_{-\sqrt{2^n}}^{\sqrt{2^n}} x^2 dF(x)-\left(\int_{-\sqrt{2^n}}^{\sqrt{2^n}} x dF(x)\right)^2 \quad \text{if } 2^n\leq k<2^{n+1},\ n=1,2\ldots.$$

A sequence of i.i.d.r.v.'s $X_1, X_2, \ldots$ with distribution function $F(x)$ and a sequence of independent normal random variables $Y_1, Y_2, \ldots$ with $\mathbf{E}Y_k=0$, $\mathbf{E}Y_k^2=\sigma_k^2$ can be constructed in such a way that the partial sums $S(n)=X_1+ +\cdots+X_n$, $T(n)=Y_1+\cdots+Y_n$, $n=1,2,\ldots$ satisfy the relation

$$|S(n)-T(n)|\overset{a.s.}{=}o(n^{1/2}).$$

We note here that in this theorem the partial sum sequence $\{S(n)\}$ is approximated by a Gaussian process $\{T(n)\}$ which is not a Wiener process. In fact $S(n)$ with two moments only cannot be approximated by a Wiener process at that rate as proved by Major [16]. He showed that for any sequence $\{a_n\}$ of real numbers with $\{a_n\} \nearrow \infty$ there exists a distribution function F with mean 0 and variance 1 such that for any i.i.d. sequence $\{X_i\}$ having the distribution F and for any Wiener process $W(t)$ we have

$$\overline{\lim_{n\to\infty}}\, a_n |S(n) - W(n)|/(n \log\log n)^{1/2} \overset{a.s.}{=} \infty.$$

In other words this result is saying that if we assume the existence of two moments only, then the rate of approximation of Strassen's theorem (cf. Strassen [22]) is the best possible if $S(n)$ is to be approximated by a Wiener process. At the same time, the rate of approximation in Theorem D is just enough to prove the weighted approximations of Theorem 2.1, as well as those of Theorems 3.2, 3.3 and 3.4 on $[1, \infty)$.

We note also that Theorem D (Major, [17]) is that strong invariance principle which implies Donsker's weak convergence theorem, as well as Strassen [22] strong invariance principle for the law of the iterated logarithm (cf., also Csörgő and Révész, [9] p. 112). By proving Theorem 2.1, we show that Theorem D is strong enough to imply also the optimal "weighted" Donker's theorem in $D[0,1]$.

PROOF OF THEOREM 2.1. Let $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ be as in Theorem D and $\{W(t),\ 0 \leqq t < \infty\}$ be a Wiener process such that

$$W(n) = \sum_{i=1}^{n} Y_i/\sigma_i, \qquad n = 1, 2 \ldots . \tag{2.7}$$

Let $T(nt) = \sum_{i=1}^{[nt]} Y_i$, $0 \leqq t \leqq 1$, and $q \in Q$. We have

$$\begin{aligned}\sup_{0<t\leqq 1} \left| n^{-1/2}(S(nt) - W(nt)) \right| / q(t) &= \sup_{0<t<1/n} n^{-1/2} |W(nt)|/q(t) \\ &+ \sup_{1/n\leqq t<1} \left| n^{-1/2}(S(nt) - T(nt)) \right| / q(t) \\ &+ \sup_{1/n\leqq t<1} \left| n^{-1/2}(T(nt) - W(nt)) \right| / q(t) \\ &= I_1(n) + I_2(n) + I_3(n).\end{aligned}$$

Since for each $n \geqq 1$

$$\left\{ n^{-1/2} W(nt)/q(t),\ 0 < t \leqq 1 \right\} \overset{\mathcal{D}}{=} \{W(t)/q(t),\ 0 < t \leqq 1\}, \tag{2.8}$$

by Theorems C and B respectively, we obtain

$$I_1(n) = \begin{cases} o_P(1) & \text{if } I(q,c) < \infty \text{ for all } c > 0 \\ O_P(1) & \text{if } I(q,c) < \infty \text{ for some } c > 0. \end{cases} \tag{2.9}$$

By Theorem D we have

$$|S(nt) - T(nt)| \overset{a.s.}{=} o((nt)^{1/2}), \quad \text{as } nt \to \infty,$$

and consequently

$$\sup_{1 \leqq nt < \infty} |S(nt) - T(nt)|/(nt)^{1/2} \overset{a.s.}{=} O(1), \quad \text{as } n \to \infty.$$

Let $\delta \in (0,1)$ be fixed and n be such that $1/n < \delta$. Then, a.s. as $n \to \infty$,

$$\begin{aligned} \sup_{1/n \leqq t < \delta} \left| n^{-1/2}(S(nt) - T(nt)) \right| /q(t) \\ \leqq O(1) \sup_{0<t<\delta} t^{1/2}/q(t). \end{aligned} \tag{2.10}$$

Using Theorem D once again, we get

$$\sup_{\delta \leqq t < 1} \left| n^{-1/2}(S(nt) - T(nt)) \right| /q(t) \overset{a.s.}{=} o(1) \tag{2.11}$$

for any $\delta \in (0,1)$. Taking $\delta > 0$ arbitrarily small, by (2.10) and (2.11) we conclude

$$I_2(n) = \sup_{1/n \leqq t \leqq 1} \left| n^{-1/2}(S(nt) - T(nt)) \right| /q(t) \overset{a.s.}{=} o(1) \tag{2.12}$$

for any $q \in Q$ such that

$$\lim_{t \downarrow 0} t^{1/2}/q(t) = 0. \tag{2.13}$$

Next we note that for each $n \geqq 1$

$$\left\{ n^{-1/2} \sum_{i=1}^{[nt]} \left(1 - \frac{1}{\sigma_i}\right) Y_i,\ 0 \leqq t \leqq 1 \right\} \overset{\mathcal{D}}{=} \left\{ W\left(\frac{1}{n} \sum_{i=1}^{[nt]} (\sigma_i - 1)^2 \right),\ 0 \leqq t \leqq 1 \right\}.$$

Let $\delta > 0$ be small enough, so that q is already nondecreasing on $(0, \delta)$ and let n be such that $1/n < \delta$. Then

$$\begin{aligned} &\sup_{1/n \leqq t < \delta} \left| n^{-1/2} \left(T(nt) - \sum_{i=1}^{[nt]} Y_i/\sigma_i \right) \right| \Big/ q(t) \\ &\overset{\mathcal{D}}{=} \sup_{1/n \leqq t < \delta} \left| W\left(\frac{1}{n} \sum_{i=1}^{[nt]} (\sigma_i - 1)^2 \right) \right| \Big/ q(t) \\ &\leqq \sup_{1/n \leqq t < \delta} \left| W\left(\frac{1}{n} \sum_{i=1}^{[nt]} (\sigma_i - 1)^2 \right) \right| \Big/ q\left(\frac{1}{n} \sum_{i=1}^{[nt]} (\sigma_i - 1)^2 \right). \end{aligned} \tag{2.14}$$

By Kolmogorov's inequality

$$\sup_{0\leqq t\leqq 1}\left|n^{-1/2}\left(T(nt)-\sum_{i=1}^{[nt]}Y_i/\sigma_i\right)\right| = o_P(1)$$

which implies

$$\sup_{\delta\leqq t\leqq 1}\left|n^{-1/2}\left(T(nt)-\sum_{i=1}^{[nt]}Y_i/\sigma_i\right)\right|\Big/q(t) = o_P(1) \tag{2.15}$$

for any $\delta\in(0,1)$. Since $\delta>0$ can be taken arbitrarily small, using again Theorems C and B, by (2.14) and (2.15) we conclude

$$\begin{aligned}&\sup_{1/n\leqq t\leqq 1}\left|n^{-1/2}\left(T(nt)-\sum_{i=1}^{[nt]}Y_i/\sigma_i\right)\right|\Big/q(t)\\ &=\begin{cases}o_P(1) & \text{if } I(q,c)<\infty \text{ for all } c>0\\ O_P(1) & \text{if } I(q,c)<\infty \text{ for some } c>0.\end{cases}\end{aligned} \tag{2.16}$$

Since

$$|W(nt)-W([nt])| = O((\log nt)^{1/2}) \quad \text{a.s., as } nt\to\infty,$$

(cf., e.g., Theorem 1.2.1 of Csörgő and Révész, [9]), by arguing similarly as for (2.12), we have

$$\sup_{1/n\leqq t\leqq 1}\left|n^{-1/2}(W([nt])-W(nt))\right|/q(t) \overset{a.s.}{=} o(1) \tag{2.17}$$

for any $q\in Q$ such that (2.13) holds. Since Theorem B implies (2.13) for any $q\in Q$ such that $I(q,c)<\infty$ for some $c>0$, combining (2.8), (2.16) and (2.17) we obtain

$$I_3(n)=\begin{cases}o_P(1) & \text{if } I(q,c)<\infty \text{ for all } c>0\\ O_P(1) & \text{if } I(q,c)<\infty \text{ for some } c>0.\end{cases} \tag{2.18}$$

Consequently (2.9), (2.12) and (2.18) yield the "if" parts of (a) and (b).

The converse parts of (a) and (b) follow from Theorems C and B respectively, since

$$\sup_{0<t\leqq 1}|n^{-1/2}(S(nt)-W(nt))|/q(t) \geqq \sup_{0<t<1/n}|n^{-1/2}W(nt)|/q(t).$$

Namely, assuming (2.5), we have for any $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbf{P}\left\{\sup_{0<t<1/n} |n^{-1/2}W(nt)|/q(t) > \varepsilon\right\}$$
$$= \lim_{n\to\infty} \mathbf{P}\left\{\sup_{0<t<1/n} |W(t)|/q(t) > \varepsilon\right\}$$
$$= 0.$$

Furthermore, since $A_n(\varepsilon) := \left\{\omega : \sup_{0<t<1/n} |W(t)|/q(t) > \varepsilon\right\}$ is a decreasing sequence, we obtain

$$\mathbf{P}\left\{\bigcap_{n=1}^{\infty} A_n(\varepsilon)\right\} = 0$$

and, consequently,

$$\mathbf{P}\{A_n(\varepsilon) \text{ i.o.}\} = 0,$$

i.e.,

$$\lim_{n\to\infty} \sup_{0<t<1/n} |W(t)|/q(t) = 0 \quad \text{a.s..}$$

Now Theorem C implies that $I(q,c) < \infty$ for all $c > 0$.

Assuming (2.6), one shows similarly that

$$\limsup_{n\to\infty} \sup_{0<t<1/n} |W(t)|/q(t) < \infty \quad \text{a.s..}$$

By Theorem B we get $I(q,c) < \infty$ for some $c > 0$.

Proof of Corollary 2.1. Let $S_n(t) = S(nt)/n^{1/2}$.

Obviously, with $q \in Q$ and $I(q,c) < \infty$ for all $c > 0$, by Theorem 2.1 (a) we have

$$S_n(t)/q(t) \xrightarrow{\mathcal{D}} W(t)/q(t) \quad \text{in } D[0,1], \tag{2.19}$$

where W is a standard Wiener process.

Assuming now (2.19) with $q \in Q$, by Skorohod–Dudley–Wichura theorem (cf. Shorack and Wellner, [21], p. 47) there exists S_n^*, $n \geqq 1$, and W^* such that

$$\left(\frac{S_n}{q}\right)^* \stackrel{\mathcal{D}}{=} \frac{S_n}{q}, \quad n \geqq 1, \qquad \left(\frac{W}{q}\right)^* \stackrel{\mathcal{D}}{=} \frac{W}{q}$$

and

$$\sup_{0<t\leqq 1} \left| \left(\frac{S_n}{q}\right)^*(t) - \left(\frac{W}{q}\right)^*(t) \right| = o(1) \quad \text{a.s..}$$

Hence

$$\sup_{0<t<1/n}\left|\left(\frac{W}{q}\right)^{*}(t)\right| = o(1) \quad \text{a.s.}$$

which, by Theorem C implies $I(q,c)<\infty$ for all $c>0$.

COROLLARY 2.2. *Let $X_1, X_2, \dots$ be i.i.d.r.v.'s such that*

$$\mathbf{E}X_1 = 0, \qquad \mathbf{E}X_1^2 = 1.$$

Then with $q \in Q$ the following three statements are equivalent:

(a) *There exists a standard Wiener process such that, as $n \to \infty$*

$$\sup_{0<t\leqq 1}\left|n^{-1/2}(S(nt)-W(nt))\right|/q(t) = o_P(1).$$

(b) *As $n\to\infty$*

$$n^{-1/2}S(nt)/q(t)\xrightarrow{\mathcal{D}}W(t)/q(t) \quad \text{in } D[0,1]$$

where W is a standard Wiener process.

(c)

$$I(q,c)<\infty \text{ for all } c>0.$$

Obviously, Corollary 2.1 implies weak convergence of any continuous in sup-norm functional of $n^{-1/2}S(nt)/q(t)$ to the corresponding functional of $W(t)/q(t)$ with $q\in Q$ and such that $I(q,c)<\infty$ for all $c>0$. However, for the sup- functional itself the class of possible weight functions is bigger. Such a phenomenon was first noticed and proved for weighted empirical and quantile processes by [2], and then for partial sums when assuming the existence of more than two moments by Csörgő and Horváth [4]. Here we improve on the latter result by assuming the existence of two moments only. First we prove the following result.

THEOREM 2.2. *Let $X_1, X_2, \dots$ be i.i.d.r.v.'s such that*

$$\mathbf{E}X_1 = 0, \qquad \mathbf{E}X_1^2 = 1.$$

Let $q\in Q$ be such that $I(q,c)<\infty$ for some $c>0$ and $q(t)/t^{1/2}$ is slowly varying at 0. Then, as $n\to\infty$, we have

$$\sup_{1/n\leqq t\leqq 1}\left|n^{-1/2}(S(nt)-W(nt))\right|/q(t) = o_P(1). \tag{2.20}$$

We note that since $\limsup_{t\downarrow 0}|W(t)|/t^{1/2}=\infty$, weight functions in our considerations are necessary of the form

$$q(t)=t^{1/2}l(t), \text{ where } l(t)\to\infty \text{ as } t\downarrow 0. \tag{2.21}$$

We recall that $l(t)$ is a slowly varying function at 0 if it is positive, measurable and

$$\lim_{t\downarrow 0}\frac{l(\epsilon t)}{l(t)}=1 \text{ for all } \epsilon>0. \tag{2.22}$$

We note also that, for example, $l(t)=(\log\log((1/t)\vee 3))^{1/2}$ is slowly varying at 0 and hence Theorem 2.2 holds with $q(t)=(t\log\log((1/t)\vee 3))^{1/2}$.

PROOF OF THEOREM 2.2. Let $X_1,X_2,\ldots$ and $Y_1,Y_2,\ldots$, with $T(k)==Y_1+\cdots+Y_k$, be as in Theorem D and $\{W(t),\ 0\leqq t<\infty\}$ be a standard Wiener process such that

$$W(n)=\sum_{i=1}^{n}Y_i/\sigma_i, \qquad n=1,2,\ldots .$$

We assume that $q\in Q$ is such that $q(t)/t^{1/2}$ is slowly varying function at 0 and $I(q,c)<\infty$ for some $c>0$. Arguing as in the proof of Theorem 2.1, we have

$$\sup_{1/n\leqq t\leqq 1}\left|n^{-1/2}(S(nt)-T(nt))\right|/q(t)=o(1), \quad \text{a.s.} \tag{2.23}$$

as well as

$$\sup_{\delta\leqq t\leqq 1}\left|n^{-1/2}\left(T(nt)-\sum_{i=1}^{[nt]}Y_i/\sigma_i\right)\right|\Big/q(t)=o_P(1) \tag{2.24}$$

for any $\delta\in(0,1)$. Let $\delta>0$ be small enough, so that q is already non-decreasing on $(0,\delta)$ and let n be such that $1/n<\delta$. We have

$$\begin{aligned}
&\sup_{1/n\leqq t<\delta}\left|n^{-1/2}\left(T(nt)-\sum_{i=1}^{[nt]}Y_i/\sigma_i\right)\right|\Big/q(t)\\
&=\sup_{1/n\leqq t\leqq q(1/n)/n^{1/2}}\left|n^{-1/2}\sum_{i=1}^{[nt]}\left(1-\frac{1}{\sigma_i}\right)Y_i\right|\Big/q(t)\\
&+\sup_{q(1/n)/n^{1/2}<t<\delta}\left|n^{-1/2}\sum_{i=1}^{[nt]}\left(1-\frac{1}{\sigma_i}\right)Y_i\right|\Big/q(t)\\
&=I_4(n)+I_5(n).
\end{aligned} \tag{2.25}$$

In order to show that $I_4(n)=o_P(1)$, we note that due to $q(t)/t^{1/2}\to\infty$ as $t\downarrow 0$ (cf. Theorem A) and $\sigma_i\to 1$ as $i\to\infty$, for any $\varepsilon>0$ there is n large enough such that

$$\frac{1}{n^{1/2}q(1/n)} \sum_{i=1}^{n^{1/2}q(1/n)} (\sigma_i - 1)^2 \leqq \varepsilon^2.$$

Consequently, and by Kolmogorov's inequality, we have

$$\mathbf{P}\left\{ \sup_{1/n \leqq t \leqq q(1/n)/n^{1/2}} \left| \sum_{i=1}^{[nt]} \left(1 - \frac{1}{\sigma_i}\right) Y_i \right| \Big/ n^{1/2} q(t) > \varepsilon \right\}$$

$$\leqq \mathbf{P}\left\{ \sup_{1/n \leqq t \leqq q(1/n)/n^{1/2}} \left| \sum_{i=1}^{[nt]} \left(1 - \frac{1}{\sigma_i}\right) Y_i \right| \Big/ n^{1/2} q(1/n) > \varepsilon \right\}$$

$$\leqq \frac{\sum_{i=1}^{n^{1/2}q(1/n)} (\sigma_i - 1)^2}{nq^2(1/n)\varepsilon^2}$$

$$\leqq \frac{\varepsilon^2}{n^{1/2}q(1/n)\varepsilon^2}$$

$$= \frac{1}{n^{1/2}q(1/n)}.$$

Since $q(t)/t^{1/2} \to \infty$ as $t \downarrow 0$, we obtain

$$I_4(n) = o_P(1) \tag{2.26}$$

as $n \to \infty$.

Next we show that $I_5(n) = o_P(1)$. Since $\sigma_i \to 1$ as $i \to \infty$, for any $\varepsilon > > 0$ there is n large enough such that $\frac{1}{nt} \sum_{i=1}^{[nt]} (\sigma_i - 1)^2 \leqq \varepsilon^2$ whenever $[nt] \geqq \geqq n^{1/2}q(1/n)$, which gives $q\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i - 1)^2\right) \leqq q(t\varepsilon^2)$. Hence, and using also (2.21), we have

$$I_5(n) \overset{\mathcal{D}}{=} \sup_{q(1/n)/n^{1/2} < t < \delta} \left| W\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i - 1)^2\right) \right| \Big/ q(t)$$

$$= \sup_{q(1/n)/n^{1/2} < t < \delta} \frac{\left| W\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i - 1)^2\right) \right|}{q\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i - 1)^2\right)} \, \frac{q\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i - 1)^2\right)}{q(t)}$$

$$\leq \sup_{q(1/n)/n^{1/2}<t<\delta} \frac{\left| W\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)\right|}{q\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)} \frac{q(t\varepsilon^2)}{q(t)}$$

$$= \sup_{q(1/n)/n^{1/2}<t<\delta} \frac{\left| W\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)\right|}{q\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)} \frac{(t\varepsilon^2)^{1/2}l(t\varepsilon^2)}{t^{1/2}l(t)}$$

$$\leq \varepsilon \sup_{q(1/n)/n^{1/2}<t<\delta} \frac{\left| W\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)\right|}{q\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)} \sup_{0<t<\delta} \frac{l(t\varepsilon^2)}{l(t)}$$

for any $\varepsilon > 0$.

Consequently, by taking $\delta \to 0$ and hence also $n \to \infty$, and using (2.2), Theorem B and (2.22), we arrive at

$$\sup_{q(1/n)/n^{1/2}<t<\delta} \left| W\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)\right| \Big/ q(t) \overset{a.s.}{=} \varepsilon O(1)(1+o(1)).$$

Since $\varepsilon > 0$ can be taken arbitrarily small, we have

$$I_5(n) = o_P(1). \tag{2.27}$$

Combining (2.23)–(2.27) we get (2.20).

THEOREM 2.3. *Let $X_1, X_2, \ldots$ be i.i.d.r.v.'s such that*

$$\mathbf{E}X_1 = 0, \qquad \mathbf{E}X_1^2 = 1,$$

and $q \in Q$ is such that $q(t)/t^{1/2}$ is slowly varying function at 0. Then, as $n \to \infty$,

$$\sup_{0<t\leq 1} \left| n^{-1/2} S(nt)\right| / q(t) \xrightarrow{\mathcal{D}} \sup_{0<t\leq 1} |W(t)|/q(t)$$

if and only if $I(q,c) < \infty$ for some $c > 0$, where W is a standard Wiener process.

In particular we have, as $n \to \infty$

$$\sup_{0<t\leqq 1} \left| n^{-1/2} S(nt) \right| \Big/ \left(t \log\log\left(\frac{1}{t} \vee 3\right)\right)^{1/2}$$
$$\xrightarrow{\mathcal{D}} \sup_{0<t\leqq 1} |W(t)| \Big/ \left(t \log\log\left(\frac{1}{t} \vee 3\right)\right)^{1/2}.$$

We note also that for the weight function $q(t) = (t \log\log((1/t) \vee 3))^{1/2}$ Theorem 2.1 (a) and Corollary 2.1 does not hold.

PROOF. Let $q \in Q$ be such that $q(t)/t^{1/2}$ is a slowly varying function at 0 and $I(q, c) < \infty$ for some $c > 0$. By Theorem 2.2 we have

$$\sup_{0<t\leqq 1} \left| n^{-1/2}(S(nt) - \overline{W}(nt)) \right| / q(t) = o_P(1)$$

as $n \to \infty$, where

$$\overline{W}(nt) = \begin{cases} W(nt) & \text{for } 1/n \leqq t \leqq 1 \\ 0 & \text{elsewhere} \end{cases}$$

and $W(\cdot)$ is a standard Wiener process. Hence

$$\lim_{n\to\infty} \left| \mathbf{P}\left\{ \sup_{0<t\leqq 1} |S(nt)|/n^{1/2} q(t) \leqq x \right\} - \mathbf{P}\left\{ \sup_{1/n\leqq t<1} |W(t)|/q(t) \leqq x \right\} \right| = 0$$

for any Wiener process and $-\infty < x < \infty$.

On account of the monotonicity of the sequence of events

$$B_n = \left\{ \omega : \sup_{1/n\leqq t<1} |W(t)|/q(t) \leqq x \right\}$$

we have

$$\lim_{n\to\infty} \mathbf{P}(B_n) = \mathbf{P}\left\{ \omega : \sup_{0<t\leqq 1} |W(t)|/q(t) \leqq x \right\}$$

for any Wiener process and $q \in Q$ such that $I(q, c) < \infty$ for some $c \geqq 0$. Consequently, we have for any $x \geqq 0$

$$\lim_{n\to\infty} \mathbf{P}\left\{ \sup_{0\leqq t\leqq 1} n^{-1/2} \frac{|S(nt)|}{q(t)} \leqq x \right\} = \mathbf{P}\left\{ \sup_{0<t\leqq 1} \frac{|W(t)|}{q(t)} \leqq x \right\}.$$

Conversely, assume that, as $n \to \infty$, with $q \in Q$ we have

$$\sup_{0<t\leqq 1} |n^{-1/2}S(nt)|/q(t) \xrightarrow{\mathcal{D}} \sup_{0<t\leqq 1} |W(t)|/q(t), \tag{2.28}$$

where $W(t)$ is a Wiener process. Hence the limiting random variable in (2.28) is almost surely finite. Consequently, (2.2) and Theorem B implies $I(q,c) < \infty$ for some $c > 0$.

The optimal conditions for weighted L_p-convergence and approximation of the empirical and quantile processes were given by Csörgő, Horváth and Shao [7]. For weighted L_p-approximations of the partial sums $S(nt)$ when only two moments are assumed to be finite, we refer to Szyszkowicz [25], [28].

3. Weighted approximations on $[1,\infty)$

Let $X_1, X_2, \dots$ be i.i.d.r.v.'s and for each $n \geqq 1$ let $S(nt) = \sum_{i=1}^{[nt]} X_i$, $0 \leqq t < \infty$. Let $\{W(t),\ 0 \leqq t < \infty\}$ be a standard Wiener process.

A function $h : [1,\infty) \to (0,\infty)$ will be called positive if $\inf_{1\leqq t\leqq K} h(t) > 0$ for all $1 < K < \infty$.

THEOREM 3.1. *Let $X_1, X_2, \dots$ be i.i.d.r.v.'s such that*

$$\mathbf{E}X_1 = 0, \quad \mathbf{E}X_1^2 = 1, \quad \mathbf{E}|X_1|^r < \infty \text{ for some } r > 2.$$

Then a standard Wiener process $\{W(t),\ 0 \leqq t < \infty\}$ can be so constructed that with a function $h(t)$ on $[1,\infty)$ which is positive and such that

$$\limsup_{t\to\infty} t^{1/2}/h(t) < \infty,$$

as $n \to \infty$, we have

$$\sup_{1\leqq t<\infty} |n^{-1/2}(S(nt) - W(nt))|/h(t) = o(1) \quad a.s..$$

PROOF. By the KMT [14] approximation one can construct a standard Wiener process $\{W(t),\ 0 \leqq t < \infty\}$ such that

$$|S(nt) - W(nt)| = o((nt)^{1/r}) \quad \text{a.s.}$$

as $nt \to \infty$. Hence

$$\sup_{1\leqq t<\infty} |n^{-1/2}(S(nt) - W(nt))|/t^{1/2} = o(1) \quad \text{a.s.}$$

as $n \to \infty$. Consequently, for any $1 < K < \infty$, we have, as $n \to \infty$

$$\sup_{K<t<\infty} |n^{-1/2}(S(nt) - W(nt))|/h(t)$$
$$\leqq o(1) \sup_{K<t<\infty} t^{1/2}/h(t) \quad \text{a.s.,}$$

while

$$\sup_{1\leqq t<K} |n^{-1/2}(S(nt) - W(nt))|/h(t) = o(1) \quad \text{a.s.,}$$

due to the positivity of h. These two statements imply the result.

From Theorem 3.1 we conclude the weak convergence of weighted partial sum processes $n^{-1/2}S(nt)/h(t)$ to $W(t)/h(t)$ whenever the limiting process is finite. Now we describe the class of weight functions for which this statement is true.

Let $\mathcal{H}$ be the class of those positive functions h on $[1,\infty)$ for which $h(t)/t$ is nonincreasing in a neighbourhood of infinity. A function $h \in \mathcal{H}$ will be called a global function of a standard Wiener process $\{W(t),\ 0 \leqq t < \infty\}$ if

$$\limsup_{t\to\infty} |W(t)|/h(t) < \infty \quad \text{a.s..}$$

Introduce the following integrals:

$$E_\infty(h,c) = \int_1^\infty t^{-3/2}h(t)\exp(-ct^{-1}h^2(t))dt$$

and

$$I_\infty(h,c) = \int_1^\infty t^{-1}\exp(-ct^{-1}h^2(t))dt,$$

where $0 < c < \infty$.

For a global description of the behaviour of a Wiener process near infinity, as well as for the following two results which are analogs of Theorem B and C for the case of $t \to \infty$, we refer to Csörgő, Shao and Szyszkowicz [10].

THEOREM B*. *A function $h \in \mathcal{H}$ is a global function of a standard Wiener process if and only if the integral $I_\infty(h,c) < \infty$ for some $c > 0$ or, equivalently, if and only if the integral $E_\infty(h,c) < \infty$ for some $c > 0$ and $\lim_{t\to\infty} h(t)/t^{1/2} = \infty$.*

THEOREM C*. *Let $h \in \mathcal{H}$ and W be a standard Wiener process. Then*

$$\lim_{t\to\infty} |W(t)|/h(t) = 0 \quad a.s.$$

if and only if the integral $I_\infty(h,c) < \infty$ for all $c > 0$ or, equivalently, if and only if the integral $E_\infty(h,c) < \infty$ for all $c > 0$ and $\lim_{t\to\infty} h(t)/t^{1/2} = \infty$.

Combining Theorems 3.1 and B*, we obtain the following result.

COROLLARY 3.1. *Let $X_1, X_2, \ldots$ be i.i.d.r.v.'s such that*

$$\mathbf{E}X_1 = 0, \quad \mathbf{E}X_1^2 = 1, \quad \mathbf{E}|X_1|^r < \infty \text{ for some } r > 2.$$

Let $h \in \mathcal{H}$ and $\{W(t),\ 0 \leqq t < \infty\}$ be a standard Wiener process. Then for all measurable, bounded, continuous functions $g\colon D[1,\infty) \to \mathbf{R}$, we have

$$g\left(n^{-1/2} S(n\cdot)/h(\cdot)\right) \xrightarrow{\mathcal{D}} g(W(\cdot)/h(\cdot)) \tag{3.1}$$

if and only if $I_\infty(h,c) < \infty$ for some $c > 0$.

We note that this result is not completely analogous to the case when $t \in [0,1]$, where the corresponding class of possible weight functions was smaller (cf. Corollary 2.1). In this regard we wish to point out that if $q \in Q$ then $q(1/t)$ is well defined for $t \in [1,\infty)$, positive and non-increasing in t as $t \to \infty$. Hence $tq(1/t) \in \mathcal{H}$, and our results on $[0,1]$ and on $[1,\infty)$ can be stated in terms of the integral $I(q,c)$ (or $E(q,c)$) for both cases (cf. also Csörgő, Shao and Szyszkowicz [10]).

REMARK 3.1. According to Theorem B* the statement (3.1) is also equivalent to $E_\infty(h,c) < \infty$ for some $c > 0$ and $\lim_{t\to\infty} h(t)/t^{1/2} = \infty$.

REMARK 3.2. Obviously, assuming the existence of the moment generating function for X_1, Theorem 3.1 would read identically and the proof would be similar. By taking for example

$$X_i = N(i) - N(i-1), \qquad i \geqq 1,$$

where $\{N(x),\ x \geqq 0\}$ is a Poisson process with intensity parameter $\lambda > 0$ and noting that

$$\sup_{k \leqq x < k+1} |N(x) - N(k)| = N(k+1) - N(k) = O(\log k) \quad \text{a.s.},$$

we can conclude Theorem 3.1 and Corollary 3.1 with

$$S(nt) \text{ replaced by } (N(nt) - nt\lambda)/\lambda^{1/2}.$$

Consequently we obtain an optimal in probability approximation of a Poisson process by Brownian motion in weighted supremum norm on $D[1,\infty)$.

The above method of proof cannot, however, be used if we assume the existence of two moments only. To handle this case, we use again the Theorem of Major [17] (cf. Theorem D) and obtain the following result.

THEOREM 3.2. *Let $X_1, X_2, \ldots$ be i.i.d.r.v.'s such that*

$$\mathbf{E}X_1 = 0, \qquad \mathbf{E}X_1^2 = 1.$$

Let $h \in \mathcal{H}$ and assume that $h(t)/t^{1/2}$ is slowly varying at ∞. Then a standard Wiener process $\{W(t),\ 0 \leqq t < \infty\}$ can be constructed in such a way that if $I_\infty(h,c) < \infty$ for some $c > 0$, then as $n \to \infty$

$$\sup_{1 \leqq t < \infty} |n^{-1/2}(S(nt) - W(nt))|/h(t) = o_P(1).$$

We note that for any positive function $h(t)$ such that

$$\limsup_{t \to \infty} |W(t)|/h(t) \overset{a.s.}{=} \beta \geqq 0$$

we have that h is of the form

$$h(t) = t^{1/2} L(t), \tag{3.2}$$

where $L(t) \to \infty$ as $t \to \infty$.

We recall that $L(t)$ is a slowly varying function at infinity if it is positive, measurable and

$$\lim_{t \to \infty} \frac{L(\varepsilon t)}{L(t)} = 1 \quad \text{for all } \varepsilon > 0. \tag{3.3}$$

We note also that, for example,

$$L(t) = (\log\log(t \vee 3))^{1/2}$$

is slowly varying at ∞ and hence Theorem 3.2 holds with $h(t) = (t \log\log(t \vee 3))^{1/2}$.

PROOF OF THEOREM 3.2. Let $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ be as in Theorem D and $\{W(t),\ 0 \leqq t < \infty\}$ be a Wiener process such that

$$W(n) = \sum_{i=1}^{n} Y_i/\sigma_i, \qquad n = 1, 2, \ldots .$$

By Theorem D, with $T(nt) = \sum_{i=1}^{[nt]} Y_i$, $0 \leqq t < \infty$, we have

$$|S(nt) - T(nt)| = o((nt)^{1/2}) \quad \text{a.s.}$$

as $nt \to \infty$. Hence

$$\sup_{1 \leqq t < \infty} |n^{-1/2}(S(nt) - T(nt))|/t^{1/2} = O(1) \quad \text{a.s.}$$

as $n \to \infty$, and for any $1 < K < \infty$, we have as $n \to \infty$

$$\sup_{1 \leqq t \leqq K} |n^{-1/2}(S(nt) - T(nt))|/h(t) = o(1) \quad \text{a.s.},$$

as well as

$$\sup_{K<t<\infty} |n^{-1/2}(S(nt)-T(nt))|/h(t) \leqq O(1) \sup_{K<t<\infty} t^{1/2}/h(t) \text{ a.s.}.$$

Consequently, taking K arbitrarily large, we obtain, as $n\to\infty$

$$\sup_{1\leqq t<\infty} |n^{-1/2}(S(nt)-T(nt))|/h(t) = o(1) \text{ a.s.} \tag{3.4}$$

for any $h\colon [1,\infty)\to(0,\infty)$ which is positive and such that $\lim_{t\to\infty} t^{1/2}/h(t)=0$. In particular for $h\in\mathcal{H}$ and such that $I_\infty(h,c)<\infty$ for some $c>0$, the latter is true.

Next we have, for any $1\leqq K<\infty$

$$\begin{aligned}
&\sup_{1<t<\infty} |n^{-1/2}(T(nt)-W(nt))|/h(t)\\
&\leqq \sup_{1\leqq t\leqq K} |n^{-1/2}(T(nt)-W(nt))|/h(t)\\
&+ \sup_{K<t<\infty} |n^{-1/2}(T(nt)-W([nt]))|/h(t)\\
&+ \sup_{K<t<\infty} |n^{-1/2}(W(nt)-W([nt]))|/h(t)\\
&= \mathcal{I}_1(n)+\mathcal{I}_2(n)+\mathcal{I}_3(n).
\end{aligned} \tag{3.5}$$

We note that for each $n\geqq 1$

$$\left\{ n^{-1/2}\sum_{i=1}^{[nt]} Y_i,\ 0\leqq t<\infty \right\} \stackrel{\mathcal{D}}{=} \left\{ W\left(\frac{1}{n}\sum_{i=1}^{[nt]}\sigma_i^2\right),\ 0\leqq t<\infty \right\}.$$

Since $\sigma_i\to 1$, using continuity of a Wiener process, we have for any positive function $h(t)\colon [1,\infty)\to(0,\infty)$, as $n\to\infty$

$$\mathcal{I}_1(n) = o_P(1). \tag{3.6}$$

We have for each $n\geqq 1$

$$\begin{aligned}
&\left\{ n^{-1/2}\left(\sum_{i=1}^{[nt]} Y_i - \sum_{i=1}^{[nt]} Y_i/\sigma_i\right),\ 0\leqq t<\infty \right\}\\
&\stackrel{\mathcal{D}}{=} \left\{ W\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right),\ 0\leqq t<\infty \right\}.
\end{aligned}$$

For any slowly varying function at infinity, $L(\cdot)$, there exists another slowly varying function at infinity, $L^*(\cdot)$, such that

$$\lim_{t\to\infty} L^*(t)/L(t) = 1 \tag{3.7}$$

and $h^*(t) = t^{1/2}L^*(t)$ is an increasing function of t (cf. Corollary 1.2.1 of de Haan [11] or Lemma 4.1 in Csörgő and Horváth [5]).

Let

$$\beta \overset{a.s.}{=} \limsup_{t\uparrow\infty} \frac{|W(t)|}{h(t)} = \limsup_{T\to\infty} \sup_{T\leqq t<\infty} \frac{|W(t)|}{h(t)}, \tag{3.8}$$

due to

$$I_\infty(h,c) < \infty \quad \text{for some } c > 0.$$

Given $\varepsilon > 0$, however small, then on account of $\sigma_i \to 1$ as $i\to\infty$, we can take K large enough so that $\frac{1}{nt}\sum_{i=1}^{[nt]}(\sigma_i - 1)^2 \leqq \varepsilon^2$ whenever $[nt] \geqq Kn$. Taking K even bigger if necessary, so that $h^*(t)$ is increasing for $n^{-1}\sum_{i=1}^{[nt]}(\sigma_i - 1)^2 < t < \infty$, where $h^*(t) = t^{1/2}L^*(t)$ and $L^*(t)$ is as in (3.7), we have

$$\sup_{K<t<\infty} \left| n^{-1/2}\left(T(nt) - \sum_{i=1}^{[nt]} Y_i/\sigma_i \right)\right| \Big/ h(t)$$

$$\overset{\mathcal{D}}{=} \sup_{K<t<\infty} \left| W\left(\sum_{i=1}^{[nt]} (\sigma_i - 1)^2/n \right)\right| \Big/ h(t)$$

$$= \sup_{K<t<\infty} \frac{\left| W\left(\sum_{i=1}^{[nt]} (\sigma_i - 1)^2/n \right)\right|}{h\left(\sum_{i=1}^{[nt]} (\sigma_i - 1)^2/n \right)} \frac{h\left(\sum_{i=1}^{[nt]} (\sigma_i - 1)^2/n \right)}{h(t)}$$

$$\leqq \sup_{K<t<\infty} \frac{\left| W\left(\sum_{i=1}^{[nt]} (\sigma_i - 1)^2/n \right)\right|}{h\left(\sum_{i=1}^{[nt]} (\sigma_i - 1)^2/n \right)} \frac{h\left(\sum_{i=1}^{[nt]} (\sigma_i - 1)^2/n \right)}{h^*\left(\sum_{i=1}^{[nt]} (\sigma_i - 1)^2/n \right)} \frac{h^*(t)}{h(t)} \frac{t^{1/2}\varepsilon}{t^{1/2}} \frac{L^*(t\varepsilon^2)}{L^*(t)}.$$

Consequently, using (3.3), (3.7) and (3.8) we have for any $\varepsilon > 0$

$$\sup_{K<t<\infty} \left| W\left(\sum_{i=1}^{[nt]} (\sigma_i - 1)^2/n \right)\right| \Big/ h(t) \overset{a.s.}{=} \beta\varepsilon O(1)$$

as $K \to \infty$.

Since $\varepsilon > 0$ can be taken arbitrarily small, we have

$$\mathcal{I}_2(n) = o_P(1). \tag{3.9}$$

On account of having, as $nt \to \infty$,

$$|W(nt) - W([nt])| = O((\log nt)^{1/2}) \quad \text{a.s.}$$

we have also

$$\sup_{1 \leqq t < \infty} |n^{-1/2}(W(nt) - W([nt]))|/t^{1/2} = o(1) \quad \text{a.s.}$$

as $n \to \infty$. Consequently, for any $1 < K < \infty$, we get as $n \to \infty$

$$\begin{aligned} &\sup_{K < t < \infty} |n^{-1/2}(W(nt) - W([nt]))|/h(t) \\ &\leqq o(1) \sup_{K < t < \infty} t^{1/2}/h(t) \quad \text{a.s.,} \end{aligned}$$

which gives

$$\mathcal{I}_3(n) = o(1) \quad \text{a.s.} \tag{3.10}$$

for any h positive and such that $\limsup t^{1/2}/h(t) < \infty$.

Combining now (3.6), (3.9) and (3.10), we obtain the result.

COROLLARY 3.2. *Let $X_1, X_2, \ldots$ be i.i.d.r.v.'s such that*

$$\mathbf{E}X_1 = 0, \qquad \mathbf{E}X_1^2 = 1.$$

Let $h \in \mathcal{H}$ and $h(t)/t^{1/2}$ be slowly varying at infinity. Then the following three statements are equivalent:

(a) *There exists a standard Wiener process $\{W(t),\ 0 \leqq t < \infty\}$ such that*

$$\sup_{1 \leqq t < \infty} |n^{-1/2}(S(nt) - W(nt))|/h(t) = o_P(1)$$

as $n \to \infty$, and

$$\sup_{1 \leqq t < \infty} |W(t)|/h(t) < \infty \quad \text{a.s..}$$

(b) *For all measurable, bounded, continuous functions $g : D[1, \infty) \to \mathbf{R}$, we have*

$$g\left(n^{-1/2}S(n\cdot)/h(\cdot)\right) \xrightarrow{\mathcal{D}} g(W(\cdot)/h(\cdot)),$$

as $n \to \infty$, where $\{W(t),\ 0 \leqq t < \infty\}$ is a standard Wiener process.

(c)

$$I_\infty(h, c) < \infty \text{ for some } c > 0.$$

Assuming $h(t)/t^{1/2}$ being slowly varying at infinity implies that, w.l.o.g., we can assume also that $h(t)$ is monotone near infinity. If we require only $h(t)$ to be non-decreasing near infinity instead of slowly varying, but so that $\lim_{t \to \infty} |W(t)|/h(t) = 0$ a.s., then we obtain the following Chibisov–O'Reilly type theorem.

THEOREM 3.3. *Let $X_1, X_2, \ldots$ be i.i.d.r.v.'s such that*

$$\mathbf{E}X_1 = 0, \qquad \mathbf{E}X_1^2 = 1.$$

Let $h \in \mathcal{H}$ and $h(t)$ be non-decreasing in a neighbourhood of infinity. Then, assuming that $I_\infty(h,c) < \infty$ for all $c > 0$, a standard Wiener process $\{W(t), 0 \leqq t < \infty\}$ can be so constructed that, as $n \to \infty$,

$$\sup_{1 \leqq t < \infty} |n^{-1/2}(S(nt) - W(nt))|/h(t) = o_P(1).$$

PROOF. Let $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ be as in Theorem D, and let $\{W(t), 0 \leqq t < \infty\}$ be a Wiener process such that

$$W(n) = \sum_{i=1}^{n} Y_i/\sigma_i, \qquad n = 1, 2, \ldots .$$

Arguing as in the proof of Theorem 3.2, with $T(nt) = \sum_{i=1}^{[nt]} Y_i$, $0 \leqq t < \infty$, we obtain, as $n \to \infty$

$$\sup_{1 \leqq t < \infty} |n^{-1/2}(S(nt) - T(nt))|/h(t) = o(1) \text{ a.s.}$$

for any $h\colon [1, \infty) \to (0, \infty)$ which is positive and such that $\lim_{t \to \infty} t^{1/2}/h(t) = 0$.

Next we have, for any $1 \leqq K < \infty$

$$\begin{aligned}
&\sup_{1 < t < \infty} |n^{-1/2}(T(nt) - W(nt))|/h(t) \\
&\leqq \sup_{1 \leqq t \leqq K} |n^{-1/2}(T(nt) - W(nt))|/h(t) \\
&+ \sup_{K < t < \infty} |n^{-1/2} T(nt)|/h(t) \\
&+ \sup_{K < t < \infty} |n^{-1/2} W(nt)|/h(t) \\
&= \mathcal{J}_1(n) + \mathcal{J}_2(n) + \mathcal{J}_3(n).
\end{aligned}$$

Arguing as for (3.6), we get as $n \to \infty$

$$\mathcal{J}_1(n) = o_P(1).$$

We have for each $n \geqq 1$

$$\left\{ n^{-1/2} \sum_{i=1}^{[nt]} Y_i,\ 0 \leqq t < \infty \right\} \stackrel{\mathcal{D}}{=} \left\{ W\left(\frac{1}{n} \sum_{i=1}^{[nt]} \sigma_i^2 \right),\ 0 \leqq t < \infty \right\}.$$

Let $n \geq 1$ be fixed and take K so large that $h(t)$ is non-decreasing for $n^{-1}\sum_{i=1}^{[nK]}\sigma_i^2 < t < \infty$. Then

$$\begin{aligned}
\mathcal{J}_2(n) &\overset{\mathcal{D}}{=} \sup_{K<t<\infty} \left| W\left(\frac{1}{n}\sum_{i=1}^{[nt]}\sigma_i^2\right)\right| \Big/ h(t) \\
&\leqq \sup_{K<t<\infty} \left| W\left(\frac{1}{n}\sum_{i=1}^{[nt]}\sigma_i^2\right)\right| \Big/ h\left(\frac{1}{n}\sum_{i=1}^{[nt]}\sigma_i^2\right) \\
&\leqq \sup_{\frac{1}{n}\sum_{i=1}^{[nK]}\sigma_i^2<t<\infty} |W(t)|/h(t).
\end{aligned}$$

Letting now $n \to \infty$ and taking K arbitrarily large, we obtain

$$\mathcal{J}_2(n) = o_P(1)$$

on account of $I_\infty(h,c) < \infty$ for all $c > 0$ and $\sigma_i \to 1$.

Since

$$\mathcal{J}_3(n) \overset{\mathcal{D}}{=} \sup_{K<t<\infty} |W(t)|/h(t),$$

by taking K arbitrarily large we obtain

$$\mathcal{J}_3(n) = o_P(1),$$

again on account of $I_\infty(h,c) < \infty$ for all $c > 0$.

COROLLARY 3.3. *Let $X_1, X_2, \dots$ be i.i.d.r.v.'s such that*

$$\mathbf{E}X_1 = 0, \qquad \mathbf{E}X_1^2 = 1.$$

Let $h \in \mathcal{H}$ and $h(t)$ be non-decreasing near infinity. Then the following three statements are equivalent:

(a) *There exists a standard Wiener process such that, as $n \to \infty$,*

$$\sup_{1\leqq t<\infty} |n^{-1/2}(S(nt) - W(nt))|/h(t) = o_P(1)$$

and

$$\lim_{t\to\infty} |W(t)|/h(t) = 0.$$

(b) *For all measurable, bounded, continuous functions $g\colon D[1,\infty) \to \mathbf{R}$ we have, as $n \to \infty$,*

$$g\left(n^{-1/2}S(n\cdot)/h(\cdot)\right) \xrightarrow{\mathcal{D}} g(W(\cdot)/h(\cdot)),$$

and

$$\lim_{t\to\infty} |W(t)|/h(t) = 0,$$

where $\{W(t),\ 0 \leqq t < \infty\}$ *is a standard Wiener process.*

(c)

$$I_\infty(h,c) < \infty \textit{ for all } c > 0.$$

Next we study the asymptotic behaviour of L_p-approximations and functionals of weighted partial sum processes for the optimal class of weight functions that is bigger than the one we obtained in the case of sup-norm approximations.

THEOREM 3.4. *Let* $X_1, X_2, \dots$ *be i.i.d.r.v.'s such that*

$$\mathbf{E}X_1 = 0, \qquad \mathbf{E}X_1^2 = 1,$$

and let $0 < p < \infty$. *Then a standard Wiener process* $\{W(t), 0 \leqq t < \infty\}$ *can be so constructed that with* $h(t)$ *on* $[1,\infty)$ *which is positive and such that*

$$\int_1^\infty t^{p/2}/h(t)dt < \infty \tag{3.11}$$

we have, as $n \to \infty$,

$$\int_1^\infty |n^{-1/2}(S(nt) - W(nt))|^p/h(t)dt = o_P(1).$$

PROOF. The proof is immediate if we also assume the existence of the moment generating function or more than two moments being finite, namely that $\mathbf{E}|X_1|^r < \infty$ for some $r > 2$. Using the KMT [14] approximation and arguing as in the proof of Theorem 3.1, one can construct a standard Wiener process such that with any $1 \leqq K < \infty$

$$\begin{aligned}
&\int_1^\infty |n^{-1/2}(S(nt) - W(nt))|^p/h(t)dt \\
&\leqq \sup_{1\leqq t\leqq K} |n^{-1/2}(S(nt) - W(nt))|^p \int_1^K 1/h(t)dt \\
&\quad + \sup_{K<t<\infty} \left(|n^{-1/2}(S(nt) - W(nt))|/t^{1/2}\right)^p \int_K^\infty t^{p/2}/h(t)dt \\
&= o(1) \quad \text{a.s.}.
\end{aligned}$$

In order to prove our result when assuming the existence of two moments only, we will use again Theorem of Major [17] (cf. Theorem D).

Let $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ be as in Theorem D and $\{W(t),\ 0 \leqq t < \infty\}$ be a Wiener process such that

$$W(n) = \sum_{i=1}^{n} Y_i/\sigma_i, \qquad n = 1, 2, \ldots .$$

Arguing as in the proof of Theorem 3.2 we obtain as $n \to \infty$

$$\begin{aligned}
&\int_1^\infty |n^{-1/2}(S(nt) - T(nt))|^p/h(t)dt \\
&\leqq \sup_{1 \leqq t \leqq K} |n^{-1/2}(S(nt) - T(nt))|^p \int_1^K 1/h(t)dt \\
&\quad + \sup_{K < t < \infty} \left(|n^{-1/2}(S(nt) - T(nt))|/t^{1/2}\right)^p \int_K^\infty t^{p/2}/h(t)dt \\
&= o(1)O(1) + O(1) \int_K^\infty t^{p/2}/h(t)dt \quad \text{a.s.},
\end{aligned}$$

for any $1 \leqq K < \infty$. Hence, by taking K arbitrarily large, we get

$$\int_1^\infty |n^{-1/2}(S(nt) - T(nt))|^p/h(t)dt = o(1) \quad \text{a.s.} \tag{3.12}$$

for any h on $[1, \infty)$ which is positive and such that (3.11) holds.

Arguing again as in the proof of Theorem 3.2, we obtain as $n \to \infty$

$$\int_1^K |n^{-1/2}(T(nt) - W(nt))|^p/h(t)dt = o_P(1) \tag{3.13}$$

for any $1 \leqq K < \infty$. Also for any $1 \leqq K < \infty$ and each $n \geqq 1$ we have

$$\begin{aligned}
&\int_K^\infty |n^{-1/2}(T(nt) - W([nt]))|^p/h(t)dt \\
&\overset{\mathcal{D}}{=} \int_K^\infty \left| W\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i - 1)^2\right)\right|^p \Big/ h(t)dt.
\end{aligned} \tag{3.14}$$

By Markov's inequality, for any $\varepsilon > 0$

$$\mathbf{P}\left\{\int_K^\infty \left| W\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i - 1)^2\right)\right|^p \Big/ h(t)dt > \varepsilon\right\} \tag{3.15}$$
$$\leqq \left(\mathbf{E}|N(0,1)|^p \int_K^\infty t^{p/2}/h(t)dt\right) \Big/ \varepsilon.$$

Due to

$$|W(nt) - W([nt])| = O((\log nt)^{1/2}) \quad \text{a.s.}$$

as $nt \to \infty$, we have for any $1 \leqq K < \infty$

$$\int_K^\infty |n^{-1/2}(W(nt) - W([nt]))|^p/h(t)dt$$
$$\leqq \sup_{K<t<\infty} \left(|n^{-1/2}(W(nt) - W([nt]))|/t^{1/2}\right)^p \int_K^\infty t^{p/2}/h(t)dt \tag{3.16}$$
$$= o(1) \int_K^\infty t^{p/2}/h(t)dt$$

as $n \to \infty$. Combining (3.13)–(3.16) and taking K arbitrarily large gives

$$\int_1^\infty |n^{-1/2}(T(nt) - W(nt))|^p/h(t)dt = o_P(1),$$

which, together with (3.12), yields the result.

COROLLARY 3.4. *Let $X_1, X_2, \ldots$ be i.i.d.r.v.'s with*

$$\mathbf{E}X_1 = 0, \qquad \mathbf{E}X_1^2 = 1,$$

$0 < p < \infty$, and h be a positive function on $[1, \infty)$. Then the following three statements are equivalent:

(a) *There exists a standard Wiener process $\{W(t),\ 0 \leqq t < \infty\}$ such that*

$$\int_1^\infty |n^{-1/2}(S(nt) - W(nt))|^p/h(t)dt = o_P(1)$$

and

$$\int_1^\infty |W(t)|^p/h(t)dt < \infty \quad a.s..$$

(b) *As* $n \to \infty$

$$\int_1^\infty |n^{-1/2}S(nt)|^p/h(t)dt \xrightarrow{\mathcal{D}} \int_1^\infty |W(t)|^p/h(t)dt.$$

(c)

$$\int_1^\infty t^{p/2}/h(t)dt < \infty.$$

Earlier we pointed out that in our weighted sup-norm results in $D[1,\infty)$ we can replace $S(nt)$ by $(N(nt)-nt\lambda)/\lambda^{1/2}$, where $\{N(x),\ x \geqq 0\}$ is a Poisson process with intensity parameter $\lambda > 0$. Naturally, here too, we can do the same in our Theorem 3.4 and Corollary 3.4.

In our proof of Corollary 3.4 we will use the following result which is an immediate corollary to Theorem 2.1 of Csörgő, Horváth and Shao [7].

COROLLARY 3.5. *Let* $\{W(t),\ 0 \leqq t < \infty\}$ *be a standard Wiener process and* $0 < p < \infty$. *Then with a positive function* h *on* $[1,\infty)$ *we have*

$$\int_1^\infty |W(t)|^p/h(t)dt < \infty \quad a.s.,$$

if and only if (3.11) *holds.*

PROOF OF COROLLARY 3.4. Since the equivalence of (a) and (c) follows from Theorem 3.4 and Corollary 3.5, it suffices to show that (b) and (c) are equivalent. First we assume that (c) holds. Obviously, for any $1 \leqq K < \infty$

$$\int_1^K |n^{-1/2}S(nt)|^p/h(t)dt \xrightarrow{\mathcal{D}} \int_1^K |W(t)|^p/h(t)dt,$$

and by Theorem 3.4

$$\int_K^\infty |n^{-1/2}S(nt)|^p/h(t)dt \leq 2^p \int_K^\infty |n^{-1/2}W(nt)|^p/h(t)dt$$
$$+2^p \int_K^\infty |n^{-1/2}(S(nt) - W(nt))|^p/h(t)dt$$
$$= 2^p \int_K^\infty |n^{-1/2}W(nt)|^p/h(t)dt + o_P(1).$$

Using now Corollary 3.5 when taking K arbitrarily large, we obtain (b). Conversely, assuming (b) to hold, the limiting random variable has to be finite. Hence Corollary 3.5 implies (c).

ACKNOWLEDGEMENTS. The author gratefully expresses her gratitude to Professor Miklós Csörgő for his valuable comments and suggestions. Thanks are also due to Rimas Norvaiša for stimulating discussions on notions of weak convergence.

NOTE added in proof. After this paper was accepted for publication we have succeeded in proving Theorems 2.2, 2.3, 3.2 and Corollary 3.2 without assuming that the appropriate weight functions are of regular variation. These details can be found in our forthcoming part II of this paper, to appear in Studia Scientiarum Mathematicarum Hungarica.

REFERENCES

[1] CHIBISOV, D., Some theorems on the limiting behaviour of an empirical distribution function, *Trudy Mat. Inst. Steklov.* **71** (1964), 104–112 (in Russian). *MR* **30** #2543. See also *Selected Transl. Math. Statist. Probab.* **6** (1964), 147–156.

[2] CSÖRGŐ, M., CSÖRGŐ, S., HORVÁTH, L. and MASON, D., Weighted empirical and quantile processes, *Ann. Probab.* **14** (1986), 31–85. *MR* **87e**:60041

[3] CSÖRGŐ, M. and HORVÁTH, L., Approximations of weighted empirical and quantile processes, *Statist. Probab. Lett.* **4** (1986), 275–280. *MR* **88m**:60095

[4] CSÖRGŐ, M. and HORVÁTH, L., Nonparametric methods for changepoints problems, *Handbook of statistics*, North-Holland, Amsterdam, 1988, 403–425. *Zbl.* **673**. 62089

[5] CSÖRGŐ, M. and HORVÁTH, L., On the distributions of the supremum of weighted quantile processes, *Studia Sci. Math. Hungar.* **25** (1990), 353–375. *MR* **92h**: 60030

[6] CSÖRGŐ, M. and HORVÁTH, L., *Weighted approximations in probability and statistics*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, Wiley, Chichester, 1993. *MR* **94c**:60060

[7] CSÖRGŐ, M., HORVÁTH, L. and SHAO, Q. M., Convergence of integrals of uniform empirical and quantile processes, *Stochastic Process. Appl.* **45** (1993), 283–294. *MR* **94f**: 60052

[8] Csörgő, M. and Révész, P., Strong approximations of the quantile process, *Ann. Statist.* **6** (1978), 882–894. *MR* **58** #18683

[9] Csörgő, M. and Révész, P., *Strong approximations in probability and statistics*, Probability and Mathematical Statistics, Academic Press, New York, 1981. *MR* **84d**:60050

[10] Csörgő, M., Shao, Q. M. and Szyszkowicz, B., A note on local and global functions of a Wiener process and some Rényi-type statistics, *Studia Sci. Math. Hungar.* **26** (1991), 239–259. *MR* **93i**:60078

[11] Haan, L. de, On regular variation and its application to the weak convergence of sample extremes, Math. Centre Tract, **32**, Math. Centrum, Amsterdam, 1975.

[12] Itô, K. and McKean, H. P. Jr., *Diffusion processes and their sample paths*, Die Grundlehren der math. Wissenschaften, Bd. 125, Academic Press, New York; Springer-Verlag, Berlin–New York, 1965. *MR* **33** #8031

[13] Komlós, J., Major, P. and Tusnády, G., An approximation of partial sums of independent R. V.'s and the sample D.F.I, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **32** (1975), 111–131. *MR* **51** #11605b

[14] Komlós, J., Major, P. and Tusnády, G., An approximation of partial sums of independent R. V.'s and the sample D. F. II, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **34** (1976), 33–58. *MR* **53** #6697

[15] Major, P., The approximation of partial sums of independent r. v.'s, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **35** (1976), 213–220. *MR* **54** #3823

[16] Major, P., The approximation of partial sums of i.i.d.r.v.'s when the summands have only two moments, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **35** (1976), 221–229. *MR* **54** #3824

[17] Major, P., An improvement of Strassen's invariance principle, *Ann. Probab.* **7** (1979), 55–61. *MR* **82f**:60069

[18] O'Reilly, N., On the weak convergence of empirical processes in sup-norm metric, *Ann. Probab.* **2** (1974), 642–651. *MR* **52** #4367

[19] Pyke, R. and Shorack, G. R., Weak convergence of a two-sample empirical process and a new approach to Chernoff–Savage theorems, *Ann. Math. Statist.* **39** (1968), 755–771. *MR* **37** #2357

[20] Rényi, A., On the theory of order statistics, *Acta. Math. Acad. Sci. Hungar.* **4** (1953), 191–231. *MR* **15**–885

[21] Shorack, G. R. and Wellner, J. A., *Empirical processes with applications to statistics*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, Wiley, New York, 1986. *MR* **88e**:60002

[22] Strassen, V., An invariance principle for the law of iterated logarithm, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **3** (1964), 211–226. *MR* **30** #5379

[23] Szyszkowicz, B., Weighted stochastic processes under contiguous alternatives, *C. R. Math. Rep. Acad. Sci. Canada* **13** (1991), 211–216. *MR* **92j**:60044

[24] Szyszkowicz, B., On $\|\cdot/q\|$-metric convergence and contiguous alternatives, Tech. Rep. Ser. Lab. Res. Stat. Prob. **191**, Carleton U. – U. Ottawa, 1992.

[25] Szyszkowicz, B., L_p-functionals of weighted partial sum processes, *C. R. Math. Rep. Acad. Sci. Canada* **14** (1992), 31–36. *MR* **93a**:60052

[26] Szyszkowicz, B., Weak convergence of stochastic processes in weighted metrics and their applications to contiguous changepoint analysis, Ph. D. Thesis, Carleton Univ., Ottawa, 1992.

[27] Szyszkowicz, B., Weighted asymptotics of partial sum processes in $D[1,\infty)$, *C. R. Mat. Rep. Acad. Sci. Canada* **14** (1992), 273–278. *MR* **93m**:60067

[28] Szyszkowicz, B., L_p-approximations of weighted partial sum processes, *Stochastic Process. Appl.* **45** (1993), 295–308. *MR* **94c**:60063

(Received August 9, 1994)

DEPARTMENT OF MATHEMATICS AND STATISTICS
4302 HERZBERG BUILDING
CARLETON UNIVERSITY
1125 COLONEL BY DRIVE
OTTAWA, ONTARIO
K1S 5B6
CANADA

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 355–359

MULTIPLE COVERING OF THE RANGE OF A RANDOM WALK ON $\mathbb{Z}$ (ON A QUESTION OF P. ERDŐS AND P. RÉVÉSZ)

B. TÓTH

Dedicated to Pál Révész on his 60th birthday

Abstract

With probability one, for any positive integer r there are infinitely many instants when all the sites in the range of a simple symmetric random walk on $\mathbb{Z}$ had been visited at least r times. This result disproves a conjecture of Erdős and Révész.

1. The problem

Let X_i, $i=0,1,2,\dots$ be a simple symmetric random walk on $\mathbb{Z}$. We denote by

$$\mathcal{R}_j = \{\, x\in\mathbb{Z} \mid \exists i\in\{0,1,\dots,j\}:\ X_i = x \,\} \tag{1}$$

the range covered by the random walk up to time j.

$$f_r(j) = \#\Big\{ x\in\mathcal{R}_j : |\{\, 0\leqq i\leqq j \mid X_i = x \,\}| = r \Big\} \tag{2}$$

respectively

$$g_r(j) = \sum_{s=1}^{r} f_s(j) \tag{3}$$

are the number of sites in the range $\mathcal{R}_j$, visited by the random walk exactly r times, respectively not more than r times.

In the present note we give a simple proof of the following

THEOREM. *For any positive integer $r\in\mathbb{N}$*

$$\mathbf{P}\Big(g_r(j) = 0 \quad \text{infinitely often} \Big) = 1. \tag{4}$$

This result disproves a conjecture formulated repeatedly by P. Erdős and P. Révész. See Question 12 in [1], Conjecture at the end of [2] or Section 11.7 in [7].

1991 *Mathematics Subject Classification.* Primary 60J15, 60J55.
Key words and phrases. Random walk, local time.

Work supported by the Hungarian National Foundation for Scientific Research Grant No. 7275.

2. Ingredients of the proof

We denote by T_n the nth return to the origin, coming from left, i.e.

$$\begin{aligned} T_0 &= 0 \\ T_{n+1} &= \min \{ i > T_n \mid X_{i-1} = -1, \; X_i = 0 \} \end{aligned} \tag{5}$$

and by $S_n(l)$, $l \in \mathbb{Z}$, half of the number of jumps across the edge $(l-1, l)$, performed before the stopping time T_n, i.e.

$$\begin{aligned} S_n(l) &= \#\{ 0 < j \leqq T_n \mid X_{j-1} = l-1, \;\; X_j = l \} \\ &= \#\{ 0 < j \leqq T_n \mid X_{j-1} = l, \;\; X_j = l-1 \}. \end{aligned} \tag{6}$$

The basic representation of the local time process $S_n(\cdot)$ as a critical branching (Galton–Watson) process is due to F. Knight, [4]:

PROPOSITION. *$S_n(l)$ and $S_n(-l)$, $l = 0, 1, 2, \ldots$ are two independent identically distributed Markov chains on $\mathbb{N}$ with*

$$S_n(0) = n \tag{7}$$

and transition probabilities

$$\mathbf{P}\Big(S_n(l+1) = j \Big| S_n(l) = i \Big) = P(i,j) = \binom{i+j-1}{j} 2^{-(i+j)}. \tag{8}$$

LEMMA 1. *Let $S_n(\cdot)$ be the Markov chain given in the Proposition and $\rho_n = \max \{ l \mid S_n(l) > 0 \}$. Then, for $\epsilon > 0$*

$$\mathbf{P}\Big(n^{1-\epsilon} < \rho_n < n^{1+\epsilon} \Big) \to 1. \tag{9}$$

PROOF. This follows from the well-known fact that $\frac{\rho_n}{n}$ converges in distribution to a random variable with absolutely continuous distribution. See [6] or Chapter 5 of [5]. □

The clue to the proof of the Theorem is the following

LEMMA 2. *Let $S_n(l)$, $l = 0, 1, 2, \ldots$ be the Markov chain given in the Proposition. For any $n, r \in \mathbb{N}$ and $n > r > 0$*

(10)

$$\mathbf{P}\Big(S_n(\cdot) \text{ does not hit the set } \{1, 2, \ldots, r\} \text{before hitting } \{0\} \Big) > \frac{1}{1 + r 2^{2r+1}}.$$

PROOF. We consider $n > r$ fixed and do not denote explicitly dependence on these parameters throughout the following proof. Let C be the set of trajectories of the Markov chain $S(l)$ stopped at the first hitting of the set $\{0, 1, \ldots, r\}$:

$$C = \bigcup_{k=1}^{\infty} \{ (s_0, s_1, \ldots, s_k) \mid s_0 = n, \;\; s_l > r \text{ for } l = 1, \ldots, k-1, \;\; s_k \leqq r \}. \tag{11}$$

We partition $\mathcal{C}$ into two disjoint sets:

$$\mathcal{A} = \{ (s_0, s_1, \dots, s_k) \in \mathcal{C} \mid s_k = 0 \} \tag{12}$$

$$\mathcal{B} = \{ (s_0, s_1, \dots, s_k) \in \mathcal{C} \mid 1 \leq s_k \leq r \}. \tag{13}$$

Denote by Q the measure induced on $\mathcal{C}$ by the Markov chain $S(\cdot)$. Clearly

$$Q(\mathcal{C}) = Q(\mathcal{A}) + Q(\mathcal{B}) = 1. \tag{14}$$

The last equality follows from the well-known fact that the critical Galton–Watson branching process considered (governed by the transition probabilities given in (8)) eventually hits $\{0\}$ with probability one.

Now define the map

$$\Theta : \mathcal{B} \to \mathcal{A} \tag{15}$$

in the following way: for $(s_0, s_1, \dots, s_{k-1}, s_k) \in \mathcal{B}$

$$\Theta(s_0, s_1, \dots, s_{k-1}, s_k) = (s_0, s_1, \dots, s_{k-1}, r+1, 0). \tag{16}$$

From (8) follows that for any $i \geq r+1 > j > 0$

$$P(i, r+1) > 2^{-r} P(i,j) \tag{17}$$

and

$$P(r+1, 0) = 2^{-(r+1)}. \tag{18}$$

Hence it is easy to check that for any $\underline{s} \in \mathcal{B}$

$$Q(\Theta \underline{s}) > 2^{-(2r+1)} Q(\underline{s}). \tag{19}$$

As Θ is an r to 1 map, it follows that

$$Q(\mathcal{A}) > r^{-1} 2^{-(2r+1)} Q(\mathcal{B}). \tag{20}$$

From (14) and (20) eventually we get

$$Q(\mathcal{A}) > \frac{1}{1 + r 2^{2r+1}} \tag{21}$$

which is equivalent to (10). □

3. Proof of the Theorem

Due to Lemma 1 we can choose a sequence of natural numbers n_t so that

$$n_{t+1}^{1-\epsilon} > n_t^{1+\epsilon} \tag{22}$$

and

$$\mathbf{P}(n_t^{1-\epsilon} < \rho_{n_t} < n_t^{1+\epsilon}) > 1 - 2^{-t}. \tag{23}$$

For $t = 1, 2, \dots$ let $X_j^{(t)}$, $j = 0, 1, 2, \dots$ be independent random walks stopped at the stopping times

$$T^{(t)} = T_{n_t}^{(t)}. \tag{24}$$

From (22) and (23), by Borel–Cantelli Lemma follows that, with probability one, there is a t_0 such that for $t > t_0$

$$\mathcal{R}_{T^{(t)}}^{(t)} \subset \mathcal{R}_{T^{(t+1)}}^{(t+1)}. \tag{25}$$

On the other hand, from Lemma 2, by inverse Borel–Cantelli, for any fixed $r > 0$ with probability one $\mathcal{R}_{T^{(t)}}^{(t)}$ is covered at least r times, for infinitely many t's.

Now define

$$t(j) = \max\{t \mid \sum_{s=1}^{t} T^{(s)} \leqq j\} \tag{26}$$

$$X_j = X_{j-\sum_{s=1}^{t(j)} T^{(s)}}^{(t(j)+1)}. \tag{27}$$

X_j, $j \in \mathbb{N}$ is a simple symmetric random walk. The Theorem directly follows from the observations made above. □

ACKNOWLEDGEMENT. I should like to express my gratitude for the kind hospitality of the Department of Mathematics, Heriot-Watt University Edinburgh, where this work was finished.

REFERENCES

[1] ERDŐS, P. and RÉVÉSZ, P., Problems and results on random walks, *Mathematical statistics and probability theory*, Vol. B (Bad Tatzmannsdorf, 1986), eds.: P. Bauer, F. Koneczny, W. Wertz, D. Reidel, Dordrecht–Boston, 1987, 59–65. *MR* **89h**:60112

[2] ERDŐS, P. and RÉVÉSZ, P., Three problems on the random walk in $\mathbb{Z}^d$, *Studia Sci. Math. Hungar.* **26** (1991), 309–320. *MR* **93k**:60171

[3] ERDŐS, P., My work with Pál Révész, *Lecture delivered at the conference dedicated to the 60th birthday of Pál Révész*, Budapest, 8 June 1994.

[4] Knight, F.B., Random walks and a sojourn density process of Brownian motion, *Trans. Amer. Math. Soc.* **109** (1963), 56–86. *MR* **27** #4286

[5] Kurtz, T., *Approximation of population processes*, CBMS-NSF Regional Conference Series in Applied Mathematics, **36**, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1981. *MR* **82j**:60160

[6] Lindvall, T., Limit theorems for some functionals of certain Galton–Watson branching processes, *Adv. Appl. Probab.* **6** (1974), 309–321. *MR* **49** #9967

[7] Révész, P., *Random walk in random and non-random environments*, World Scientific, Singapore, 1990. *MR* **92c**:60096

(Received September 1, 1994)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

balint@math-inst.hu

Typeset by TypoTEX Ltd., Budapest
PRINTED IN HUNGARY
Akadémiai Kiadó & Nyomda, Budapest

CONTENTS

HU ISSN 0081-6906

Studia Scientiarum Mathematicarum Hungarica

EDITOR-IN-CHIEF

D. SZÁSZ

VOLUME 31
NUMBER 4
1996

AKADÉMIAI KIADÓ, BUDAPEST

STUDIA SCIENTIARUM MATHEMATICARUM HUNGARICA

A QUARTERLY OF THE HUNGARIAN ACADEMY OF SCIENCES

Studia Scientiarum Mathematicarum Hungarica publishes original papers on mathematics mainly in English, but also in German, French and Russian. It is published in yearly volumes of four issues (mostly double numbers published semiannually) by

AKADÉMIAI KIADÓ
Publishing House of the Hungarian Academy of Sciences
H–1117 Budapest, Prielle Kornélia u. 19–35

Manuscripts and editorial correspondence should be addressed to

J. Merza
Managing Editor

P.O. Box 127
H–1364 Budapest

Tel.: +36 1 118 2875   Fax: +36 1 117 7166
e-mail: merza @ math–inst.hu

Subscription information

Orders should be addressed to

AKADÉMIAI KIADÓ
P.O.Box 245
H–1519 Budapest

For 1996 volumes 31–32 are scheduled for publication. The subscription price is \$ 104.00 each, air delivery plus \$ 20.00 per volume.

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 361–374

CANONICAL PARTITION RELATIONS FOR ASCENDING FAMILIES OF FINITE SETS

H. LEFMANN

Abstract

In this paper we prove a canonical partition result concerning a multidimensional version of Hindman's theorem on finite sums due to Milliken and Taylor.

1. Introduction and statement of results

Let ω be the set of nonnegative integers. Small latin letters usually denote nonnegative integers.

For sets X let

$\mathcal{P}(X)$ be the collection of all subsets of X;

$[X]^{<\omega}$ be the collection of all finite subsets of X;

$[X]^k$ be the collection of all k-element subsets of X.

In 1930 Ramsey proved his famous pigeon hole principle for finite sets:

THEOREM 1.1 [6]. *Let c, k be nonnegative integers. Then for every partition $\Delta\colon [\omega]^k \to \{0,\dots,c-1\}$ there exists an infinite subset $X \subseteqq \omega$ such that the restriction $\Delta \restriction [X]^k$ is a constant mapping.* □

In Ramsey's theorem partitions of $[\omega]^k$ into finitely many pieces are considered. Another situation occurs if we look at arbitrary partitions $\Delta{:}[\omega]^k \to \omega$. This is the so-called canonical situation. The aim is to find an infinite subset $X \subseteqq \omega$ such that Δ acts nice on $[X]^k$. This case was settled by Erdős and Rado, who proved the following canonical version of Ramsey's theorem:

THEOREM 1.2 [1]. *Let k be a nonnegative integer. Then for every mapping $\Delta\colon [\omega]^k \to \omega$ there exist an infinite subset $X \subseteqq \omega$ and a subset $J \subseteqq \{0,\dots,k-1\}$ such that for all $\{x_0,\dots,x_{k-1}\}$, $\{x_0^*,\dots,x_{k-1}^*\} \in [X]^k$ with $x_0 < \dots < x_{k-1}$ and $x_0^* < \dots < x_{k-1}^*$ it is valid:*

$$\Delta(\{x_0,\dots,x_{k-1}\}) = \Delta(\{x_0^*,\dots,x_{k-1}^*\})$$

iff

$$\{x_j \mid j \in J\} = \{x_j^* \mid j \in J\}.$$

1991 *Mathematics Subject Classification*. Primary 05A99.
Key words and phrases. Ramsey theory, finite sums, canonical partitions.

Every subset $J \subseteqq \{0, \dots, k-1\}$ determines a type of mappings. None of these 2^k different types can be omitted without violating the theorem. It can be seen easily that, if the range of Δ is finite, necessarily J is the empty set; this is just Ramsey's theorem.

Hindman proved that, if the set of nonnegative integers is partitioned into finitely many pieces, then there exists an infinite subset $X \subseteqq \omega$ such that all finite sums from X without repetition lie in the same piece. Here it will be more convenient to work with finite sets instead of nonnegative integers.

THEOREM 1.3 [2]. *Let c be a nonnegative integer. Then for every mapping $\Delta\colon [\omega]^{<\omega} \to \{0, \dots, c-1\}$ there exist infinitely many pairwise disjoint sets $S_0, S_1 \dots \in [\omega]^{<\omega}$ such that* $\Delta \restriction \{\bigcup_{i\in I} S_i \mid I \subseteqq \omega,\ I \text{ finite and nonempty}\} =$ const. □

For the canonical case of Hindman's theorem Taylor showed that five patterns can occur:

THEOREM 1.4 [7]. *For every mapping $\Delta\colon [\omega]^{<\omega} \to \omega$ there exist infinitely many pairwise disjoint sets $S_0, S_1, \dots \in [\omega]^{<\omega}$ such that one of the following cases holds for all finite nonempty subsets I, J of ω:*

(i) $\Delta(\bigcup_{i\in I} S_i) = \Delta(\bigcup_{j\in J} S_j)$;

(ii) $\Delta(\bigcup_{i\in I} S_i) = \Delta(\bigcup_{j\in J} S_j)$ *iff* $\min I = \min J$;

(iii) $\Delta(\bigcup_{i\in I} S_i) = \Delta(\bigcup_{j\in J} S_j)$ *iff* $\max I = \max J$;

(iv) $\Delta(\bigcup_{i\in I} S_i) = \Delta(\bigcup_{j\in J} S_j)$ *iff* $\max I = \min J$ *and* $\max I = \max J$;

(v) $\Delta(\bigcup_{i\in I} S_i) = \Delta(\bigcup_{j\in J} S_j)$ *iff* $I = J$. □

None of these five cases can be omitted without violating the theorem as, besides the constant mapping, the following mappings $f_i\colon [\omega]^{<\omega} \to [\omega]^{<\omega}$, $i \in \{\min, \max, \text{min-max}, \text{id}\}$, show:

$$\begin{aligned} f_{\min}(S) &= \{\min S\} \\ f_{\max}(S) &= \{\max S\} \\ f_{\text{min-max}}(S) &= \{\min S, \max S\} \\ f_{\text{id}}(S) &= S. \end{aligned}$$

For subsets $\mathcal{A} \subseteqq [\omega]^{<\omega}$ with $\mathcal{A} = \{A_j \mid j \in J\}$ we denote by $FU(\mathcal{A})$ the set $\{\bigcup_{i\in I} A_i \mid I \subseteqq J,\ I \neq \emptyset,\ I \text{ finite}\}$ of all finite nonempty unions of elements of $\mathcal{A}$.

The usual total order $\leqq$ on ω induces a partial order $\leqq'$ on $[\omega]^{<\omega}$ by: $S <' T$ iff $\max S < \min T$. By abuse of language we use the same symbol $<$ for both orders. For subsets $I \subseteqq \omega$ we denote by $\{A_i \mid i \in I\}_<$ ascending (w.r.t. $<$) families $(A_i \mid i \in I)$ with elements from $[\omega]^{<\omega}$, i.e. $A_i < A_j$ for all $i, j \in I$ with $i < j$.

An infinite subset $\mathcal{A} \subseteq [\omega]^{<\omega}$ is an *ascending collection* iff for each two different sets $S, T \in \mathcal{A}$ it is valid: $S < T$ or $T < S$. Notice that for an ascending collection $\mathcal{A}$, possibly after some renumbering, we have $\mathcal{A} = \{A_0, A_1, \dots\}_<$.

For an ascending collection $\mathcal{A}$ let $[FU(\mathcal{A})]^k_<$ denote the set

$$\{\{S_0, \dots, S_{k-1}\}_< \mid S_0, \dots, S_{k-1} \in FU(\mathcal{A})\}$$

of k-term ascending families with elements from $FU(\mathcal{A})$.

Milliken and Taylor proved the following multidimensional version of Hindman's theorem, which also generalizes Ramsey's theorem.

THEOREM 1.5 [4], [7]. *Let c, k be positive integers and let $\mathcal{A}$ be an ascending collection. Then for every partition Δ:$[FU(\mathcal{A})]^k_< \to \{0, \dots, c-1\}$ there exists an ascending collection $\mathcal{B} \subset FU(\mathcal{A})$ such that $\Delta \restriction [FU(\mathcal{B})]^k_<$=const.* □

Prömel and Voigt asked for a canonical version of this result [5]. Since Theorem 1.5 is a combination of Ramsey's theorem and Hindman's theorem, one might expect that a canonical version is a combination of the theorem of Erdős and Rado and the theorem of Taylor. The patterns will be determined by subsets $J \subseteq \{0, \dots, k-1\}$, where each element of J is associated to the minimum-, maximum-, minimum and maximum- or identity-mapping. However, additional patterns can occur, since the entries of different components can be merged without leaving the underlying category, contrary to the Erdős–Rado case. Before we give a canonical version of Theorem 1.5, we need for technical reasons (to avoid redundancies) the following

DEFINITION. Let k_0, k be nonnegative integers with $k_0 \leqq k$. Further let $(T_i)_{i<k_0}$ be a sequence of nonempty sets and let $(f_i)_{i<k_0}$ be a sequence of mappings from $\{f_{\min}, f_{\max}, f_{\text{min-max}}, f_{\text{id}}\}$. The sequence $((T_i, f_i))_{i<k_0}$ of pairs is a *k-canonical sequence* iff $\bigcup_{i \in k_0} T_i \subseteq \{0, \dots, k-1\}$, and $T_0 < \dots < T_{k_0-1}$, and in addition for every $i < k_0 - 1$ and every $j < k_0$ it is valid:

(i) $f_i = f_{\min}$ implies $f_{i+1} \neq f_{\max}$,

(ii) $|T_j| > 1$ implies $f_j = f_{\text{min-max}}$ or $f_j = f_{\text{id}}$,

(iii) $|T_j| > 2$ implies $f_j = f_{\text{id}}$.

THEOREM 1.6. *Let k be a nonnegative integer and let $\mathcal{A}$ be an ascending collection. Then for every mapping Δ: $[FU(\mathcal{A})]^k_< \to \omega$ there exist an ascending collection $\mathcal{B} \subset FU(\mathcal{A})$ and a k-canonical sequence $((T_i, f_i))_{i<k_0}$ with the following property:*

For every $\{S_0, \dots, S_{k-1}\}_<$, $\{S^*_0, \dots, S^*_{k-1}\}_< \in [FU(\mathcal{B})]^k_<$

$$\Delta(\{S_0, \dots, S_{k-1}\}_<) = \Delta(\{S^*_0, \dots, S^*_{k-1}\}_<)$$

iff

$$\{f_i(\bigcup_{j \in T_i} S_j) \mid i < k_0\} = \{f_i(\bigcup_{j \in T_i} S^*_j) \mid i < k_0\}. \qquad \square$$

None of these patterns can be omitted without violating the theorem as it can be seen for k-canonical sequences $((T_i, f_i))_{i<k_0}$ with the mappings

$$\Delta_{((T_i,f_i))_{i<k_0}} : [FU(\mathcal{A})]^k_< \to [FU(\{\{0\},\{1\},\dots)]^{k_0}_<$$

defined by the relation

$$\Delta_{((T_i,f_i))_{i<k_0}}(\{S_0,\dots,S_{k-1}\}_<) = \{f_i(\bigcup_{j\in T_i} S_j) \mid i<k_0\}.$$

For $k=1$ Theorem 1.6 is just Taylor's result (1.4). The empty sequence describes the constant case. For $k=2$ there are 26 canonical cases. Besides the 25 cases, which arise from associating one of the five types occurring in Theorem 1.4 to the components, the two components can also be merged, i.e. $k_0=1$, $T_0=\{0,1\}$ and $f_0=f_{\mathrm{id}}$. Denoting by S_k the number of canonical cases for k-term ascending families, we get by recursion:

$$S_k = \frac{1}{13\cdot 2^{k+1}}\left((13+3\sqrt{13})(7+\sqrt{13})^k + (13-3\sqrt{13})(7-\sqrt{13})^k\right).$$

From Theorem 1.6 we obtain as a corollary the result of Erdős and Rado (1.2), since every mapping $\Delta\colon [\omega]^k\to\omega$ induces for an arbitrary ascending collection $\mathcal{A}$ another mapping $\Delta^*\colon [FU(\mathcal{A})]^k_<\to\omega$ by $\Delta^*(\{S_0,\dots,S_{k-1}\}_<)= \Delta(\{\min S_0,\dots,\min S_{k-1}\})$.

So far we have considered mappings $\Delta\colon [FU(\mathcal{A})]^k_<\to\omega$ w.r.t. one level only. In order to handle the behaviour of mappings on arbitrary structures w.r.t. different levels the concept of diversification was introduced in [8]. In the context here this means, roughly speaking, that for every pair $\Delta_0\colon [FU(\mathcal{A})]^{k_0}_<\to\omega$ and $\Delta_1\colon [FU(\mathcal{A})]^{k_1}_<\to\omega$ of mappings there exists an ascending collection $\mathcal{B}\subset FU(\mathcal{A})$ such that the restricted mappings $\Delta_0\restriction[FU(\mathcal{B})]^{k_0}_<$ and $\Delta_1\restriction[FU(\mathcal{B})]^{k_1}_<$ either are identical or have disjoint images. Using the concept of diversification we can characterize the behaviour of mappings w.r.t. different levels in the following way.

THEOREM 1.7. *Let m be a positive integer and let $\mathcal{A}$ be an ascending collection. Then for every mapping $\Delta\colon \bigcup_{k<m}[FU(\mathcal{A})]^k_<\to\omega$ there exists an ascending collection $\mathcal{B}\subseteq FU(\mathcal{A})$ and for every $k<m$ there exists a k-canonical sequence $((T_{k,j},f_{k,j}))_{j<m(k)}$, such that for every pair (k,l) with $k,l<m$ one of the following cases holds:*

(i) $\Delta([FU(\mathcal{B})]^k_<)\cap\Delta([FU(\mathcal{B})]^l_<)=\emptyset$;

(ii) *For every* $\{S_0,\dots,S_{k-1}\}_<\in[FU(\mathcal{B})]^k_<$ *and* $\{S^*_0,\dots,S^*_{l-1}\}_<\in [FU(\mathcal{B})]^l_<$:

$$\Delta(\{S_0,\dots,S_{k-1}\}_<)=\Delta(\{S^*_0,\dots,S^*_{l-1}\}_<)$$

iff

$$\{f_{k,j}(\bigcup_{i\in T_{k,j}} S_i)\mid j<m(k)\} = \{f_{l,j}(\bigcup_{i\in T_{l,j}} S^*_i)\mid j<m(l)\}. \qquad \square$$

This leads immediately to the following result of Voigt concerning partitions of the set of subsets of ω having cardinality less than some m:

COROLLARY 1.8 [8]. *Let m be a positive integer. Then for every mapping $\Delta\colon \bigcup_{k<m} [\omega]^k \to \omega$ there exists an infinite subset $X \subseteq \omega$ and for every $k < m$ there exists a subset $J_k \subseteq \{0, \dots, k-1\}$ such that for every pair (k, l) with $k, l < m$ one of the following cases is valid:*

(i) $\Delta([X]^k) \cap \Delta([X]^l) = \emptyset$;

(ii) *For every* $\{x_0, \dots, x_{k-1}\}_< \in [X]^k$ *and* $\{x_0^*, \dots, x_{l-1}^*\}_< \in [X]^l$

$$\Delta(\{x_0, \dots, x_{k-1}\}) = \Delta(\{x_0^*, \dots, x_{l-1}^*\})$$

iff

$$\{x_j \mid j \in J_k\} = \{x_j^* \mid j \in J_l\}. \qquad \square$$

In the next section we will prove Theorems 1.6 and 1.7.

2. Proofs

DEFINITION. Let $\mathcal{A}$ be an ascending collection and let $\Delta\colon [FU(\mathcal{A})]^k_< \to \omega$ be a mapping. Let $((T_i, f_i))_{i<k_0}$ be a k-canonical sequence. The mapping Δ has type $((T_i, f_i))_{i<k_0}$, type $(\Delta) = ((T_i, f_i))_{i<k_0}$, iff for every $\{S_0, \dots, S_{k-1}\}_<$, $\{S_0^*, \dots, S_{k-1}^*\}_< \in [FU(\mathcal{A})]^k_<$ we have

$$\Delta(\{S_0, \dots, S_{k-1}\}_<) = \Delta(\{S_0^*, \dots, S_{k-1}^*\}_<)$$

iff

$$\{f_i(\bigcup_{j \in T_i} S_j) \mid i < k_0\} = \{f_i(\bigcup_{j \in T_i} S_j^*) \mid i < k_0\}.$$

DEFINITION. Let $\mathcal{A}$ be an ascending collection and let $\Delta\colon [FU(\mathcal{A})]^k_< \to \omega$ be a mapping. The mapping Δ has *full type* iff there exists a k-canonical sequence $((\{i\}, f_i))_{i<k}$ such that type $(\Delta) = ((\{i\}, f_i))_{i<k}$.

If a mapping $\Delta\colon [FU(\mathcal{A})]^k \to \omega$, where $k \geqq 2$, has full type, then the entries of different components cannot be shifted without changing the image.

LEMMA 2.1. *Let $k \geqq 2$ be an integer and let $\mathcal{A}$ be an ascending collection. Let $\Delta\colon [FU(\mathcal{A})]^k_< \to \omega$ be a mapping, which has full type. Then for every $i \leqq k-2$ and for every $\{S_0, \dots, S_k\}_< \in [FU(\mathcal{A})]^{k+1}_<$ it holds:*

$$\Delta(\{S_0, \dots, S_{i-1}, S_i \cup S_{i+1}, S_{i+2}, \dots, S_k\}_<)$$
$$\neq \Delta(\{S_0, \dots, S_{i-1}, S_i, S_{i+1} \cup S_{i+2}, S_{i+3}, \dots, S_k\}_<). \qquad \square$$

This can be seen easily from the fact that

$$\Delta(\{S_0,\dots,S_{i-1},S_i\cup S_{i+1},S_{i+2},\dots,S_k\}_<)$$
$$=\Delta(\{S_0,\dots,S_{i-1},S_i,S_{i+1}\cup S_{i+2},S_{i+3},\dots,S_k\}_<)$$

implies $f_i=f_{\min}$ and $f_{i+1}=f_{\max}$, which is a contradiction since Δ has full type.

The concept of diversification is a main tool in the proof of Theorem 1.6. We show that ascending families of finite sets have the diversification property w.r.t. mappings of full type.

LEMMA 2.2. *Let k_0,k_1 be nonnegative integers and let $\mathcal{A}$ be an ascending collection. Further let $\Delta_i\colon [FU(\mathcal{A})]^{k_i}_< \to\omega$, $i=0,1$, be mappings of full type. Then there exists an ascending collection $\mathcal{B}\subset FU(\mathcal{A})$ such that one of the following cases is valid:*

(i) $\Delta_0([FU(\mathcal{B})]^{k_0}_<)\cap\Delta_1([FU(\mathcal{B})]^{k_1}_<)=\emptyset$;

(ii) $k_0=k_1$ *and* $\Delta_0(\{S_0,\dots,S_{k_0-1}\}_<)=\Delta_1(\{S_0,\dots,S_{k_0-1}\}_<)$

for every $\{S_0,\dots,S_{k_0-1}\}_<\in[FU(\mathcal{B})]^{k_0}_<$.

PROOF of Lemma 2.2. We prove the lemma by induction on k_0+k_1. The cases $k_0=0$ or $k_1=0$ are trivial. Thus let k_0,k_1 be positive integers and assume that the assumption is valid for every ascending collection $\mathcal{A}$ and each two mappings $\Delta\colon[FU(\mathcal{A})]^{k}_<\to\omega$ and $\Delta^*\colon[FU(\mathcal{A})]^{k^*}_<\to\omega$ with $k+k^*<k_0+k_1$. Let $\mathcal{A}$ be an ascending collection and let $\Delta_i\colon[FU(\mathcal{A})]^{k_i}_<\to\omega$, $i=0,1$, be mappings of full type, where w.l.o.g. $k_0\leqq k_1$. Since the mappings Δ_i, $i=0,1$, have full type, there are k_i-canonical sequences $((\{j\},f_{i,j}))_{j<k_i}$ such that type $(\Delta_i)=((\{j\},f_{i,j}))_{j<k_i}$. In the following we distinguish between the two cases $k_0<k_1$ and $k_0=k_1$.

Case A. $k_0<k_1$.

We will show the existence of an ascending collection $\mathcal{B}\subseteqq FU(\mathcal{A})$ such that $\Delta_0([FU(\mathcal{B})]^{k_0}_<)\cap\Delta_1([FU(\mathcal{B})]^{k_1}_<)=\emptyset$. Put $B_{-1}:=\min\mathcal{A}$ and $\mathcal{B}^{(-1)}:=\mathcal{A}\setminus\{\min\mathcal{A}\}$. We construct a sequence $(B_i)_{i<\omega}$ of finite nonempty sets and a sequence $(\mathcal{B}^{(i)})_{i<\omega}$ of ascending collections with the following properties: For every $i<\omega$:

(2.2.1) $$B_i\in FU(\mathcal{B}^{(i-1)})\quad\text{and}\quad\mathcal{B}^{(i)}\subseteqq FU(\mathcal{B}^{(i-1)}),$$

(2.2.2) $$B_{i-1}<B_i\quad\text{and}\quad B_i<\min\mathcal{B}^{(i)},$$

(2.2.3) $$\Delta_0(\{S_0,\dots,S_{k_0-1}\}_<)\neq\Delta_1(\{\bigcup_{j\in J}B_j\cup B_i,S^*_1,\dots,S^*_{k_1-1}\}_<)$$

for every $J\subseteqq\{0,\dots,i-1\}$ and for every
$\{S_0,\dots,S_{k_0-1}\}_<\in[FU(\{B_0,\dots,B_i\}\cup\mathcal{B}^{(i)})]^{k_0}_<$ and
$\{S^*_1,\dots,S^*_{k_1-1}\}_<\in[FU(\mathcal{B}^{(i)})]^{k_1-1}_<$.

Then for $\mathcal{B} := \{B_0, B_1, \dots\}_<$ we have

$$\Delta_0([FU(\mathcal{B})]^{k_0}_<) \cap \Delta_1([FU(\mathcal{B})]^{k_1}_<) = \emptyset.$$

Now we construct the sequences $(B_i)_{i<\omega}$ and $(\mathcal{B}^{(i)})_{i<\omega}$. Suppose that for some n, finite nonempty sets B_i and ascending collections $\mathcal{B}^{(i)}$ are constructed such that (2.2.1) up to (2.2.3) are valid for every $i < n$.

Step n: For every pair (I, J) with $I, J \subseteqq \{0, \dots, n-1\}$ let

$$\Delta^{(i)}_{(I,J)} : [FU(\mathcal{B}^{(n-1)})]^{k_1}_< \to \{0,1\}, \quad i = 0, 1,$$

be mappings defined by

$$\Delta^{(0)}_{(I,J)}(\{S_0, \dots, S_{k_1-1}\}_<) = 0$$

iff

$$\Delta_0(\{\bigcup_{i\in I} B_i \cup S_0 \cup S_1, S_2, \dots, S_{k_0}\}_<) = \Delta_1(\{\bigcup_{j\in J} B_j \cup S_0, S_1, \dots, S_{k_1-1}\}_<)$$

and

$$\Delta^{(1)}_{(I,J)}(\{S_0, \dots, S_{k_1-1}\}_<) = 0$$

iff

$$\Delta_0(\{\bigcup_{i\in I} B_i \cup S_1, S_2, \dots, S_{k_0}\}_<) = \Delta_1(\{\bigcup_{i\in J} B_j \cup S_0, S_1, \dots, S_{k_1-1}\}_<).$$

Then 2^{2n+1} applying of Theorem 1.5 to every $\Delta^{(i)}_{(I,J)}$ yields an ascending collection $\mathcal{B}^{(n-1)}_0 \subseteqq FU(\mathcal{B}^{(n-1)})$ and for every $i < 2$ and every pair (I, J) with $I, J \subseteqq \{0, \dots, n-1\}$ numbers $c^{(i)}_{(I,J)} < 2$ such that $\Delta^{(i)}_{(I,J)}([FU(\mathcal{B}^{(n-1)}_0)]^{k_1}_<) = \{c^{(i)}_{(I,J)}\}$. By Lemma 2.1 and (2.2.3) it follows that $c^{(0)}_{(I,J)} = c^{(1)}_{(I,J)} = 1$ for every (I, J). Putting $B_n := \min \mathcal{B}^{(n-1)}_0$ and $\mathcal{B}^{(n-1)}_1 := \mathcal{B}^{(n-1)}_0 \setminus \{\min \mathcal{B}^{(n-1)}_0\}$ we have

(2.2.4) For every pair (I, J) with $I, J \subseteqq \{0, \dots, n-1\}$ and for every $\{S_1, \dots, S_{k_1-1}\}_< \in [FU(\mathcal{B}^{(n-1)}_0)]^{k_1}_<$

$$\Delta_0(\{\bigcup_{i\in I} B_i \cup B_n \cup S_1, S_2, \dots, S_{k_0}\}_<) \neq \Delta_1(\{\bigcup_{j\in J} B_j \cup B_n, S_1, \dots, S_{k_1-1}\}_<) \text{ and}$$
$$\Delta_0(\{\bigcup_{i\in I} B_i \cup S_1, S_2, \dots, S_{k_0}\}_<) \neq \Delta(\{\bigcup_{j\in J} B_j \cup B_n, S_1, \dots, S_{k_1-1}\}_<).$$

Condition (2.2.4) will guarantee that certain mappings, which will be specified later, have different images on the same arguments.

Next we construct an ascending collection $\mathcal{B}^{(n)} \subset FU(\mathcal{B}_1^{(n-1)})$ such that (2.2.3) is satisfied for $i = n$. For every subset $J \subseteqq \{0, \ldots, n-1\}$, let $\Delta_1^J : [FU(\mathcal{B}_1^{(n-1)})]_<^{k_1-1} \to \omega$ be a mapping defined by

$$\Delta_1^J(\{S_1, \ldots, S_{k_1-1}\}_<) = \Delta_1(\{\bigcup_{j\in J} B_j \cup B_n, S_1, \ldots, S_{k_1-1}\}_<).$$

The mappings Δ_1^J, $J \subseteqq \{0, \ldots, n-1\}$, have full type, since Δ_1 has full type.

For ascending families $\mathcal{T} \in \bigcup_{t=0}^{k_0} [FU(\{B_0, \ldots, B_n\})]_<^t$ with $\mathcal{T} = \{T_0, \ldots, T_{t-1}\}_<$, and $\mathcal{T}' \in \bigcup_{t'=1}^{k_0} [FU(\{B_0, \ldots, B_n\})]_<^{t'}$ with $\mathcal{T}' = \{T_0', \ldots, T_{t'-1}\}_<$, let $\hat{\Delta}_0^{\mathcal{T}} : [FU(\mathcal{B}_1^{(n-1)})]_<^{k_0-t} \to \omega$ and $\check{\Delta}_0^{\mathcal{T}'} : [FU(\mathcal{B}_1^{(n-1)})]_<^{k_0-t'+1} \to \omega$ be mappings defined by

$$\hat{\Delta}_0^{\mathcal{T}}(\{S_t, \ldots, S_{k_0-1}\}_<) = \Delta_0(\{T_0, \ldots, T_{t-1}, S_t, \ldots, S_{k_0-1}\}_<)$$

and

$$\check{\Delta}_0^{\mathcal{T}'}(\{S_{t'-1}, \ldots, S_{k_0-1}\}_<) = \Delta_0(\{T_0', \ldots, T_{t'-1}' \cup S_{t'-1}, S_{t'}, \ldots, S_{k_0-1}\}_<).$$

Every mapping $\hat{\Delta}_0^{\mathcal{T}}$ has full type, also every $\check{\Delta}_0^{\mathcal{T}}$, possibly after omitting the first component.

Several applications of the induction hypothesis to the mappings Δ_1^J and $\hat{\Delta}_0^{\mathcal{T}}$, resp. Δ_1^J and $\check{\Delta}_0^{\mathcal{T}'}$ and respecting (2.2.4) yields an ascending collection $\mathcal{B}^{(n-1)} \subset FU(\mathcal{B}_1^{(n-1)})$ such that for every $J \subseteqq \{0, \ldots, n-1\}$ and every $\mathcal{T}, \mathcal{T}' \in \bigcup_{i=0}^{k_0} [FU(\{B_0, \ldots, B_n\})]_<^i$ with $|\mathcal{T}'| > 0$ we have

$$\Delta_1^J \left([FU(\mathcal{B}^{(n-1)})]_<^{k_1-1} \right) \cap \hat{\Delta}_0^{\mathcal{T}} \left([FU(\mathcal{B}^{(n-1)})]_<^{k_0-|\mathcal{T}|} \right) = \emptyset$$

and

$$\Delta_1^J \left([FU(\mathcal{B}^{(n-1)})]_<^{k_1-1} \right) \cap \check{\Delta}_0^{\mathcal{T}} \left([FU(\mathcal{B}^{(n-1)})]_<^{k_0-|\mathcal{T}'|+1} \right) = \emptyset.$$

Thus, we have

$$\Delta_0(\{S_0, \ldots, S_{k_0-1}\}_<) \neq \Delta_1(\{\bigcup_{j\in J} B_j \cup B_n, S_1^*, \ldots, S_{k_1-1}^*\}_<)$$

for every $J \subseteqq \{0, \ldots, n-1\}$ and for every $\{S_0, \ldots, S_{k_0-1}\}_< \in [FU(\{B_0, \ldots, B_n\} \cup \mathcal{B}^{(n)})]_<^{k_0}$ and $\{S_1^*, \ldots, S_{k_1-1}^*\}_< \in [FU(\mathcal{B}^{(n)})]_<^{k_1-1}$, and (2.2.3) is satisfied for $i = n$.

Finally we get a sequence $(B_i)_{i<\omega}$ of finite nonempty sets and a sequence $(\mathcal{B}^{(i)})_{i<\omega}$ of ascending collections, such that (2.2.1) up to (2.2.3) are valid for every $i<\omega$.

Case B. $k=k_0=k_1$.

Let $\Delta^*: [FU(\mathcal{A})]^k_< \to \{0,1\}$ be a mapping defined by $\Delta^*(\{S_0,\dots,S_{k-1}\}_<)=0$ iff $\Delta_0(\{S_0,\dots,S_{k-1}\}_<)=\Delta_1(\{S_0,\dots,S_{k-1}\}_<)$. By Theorem 1.5 there is an ascending collection $\mathcal{A}_0 \subseteqq FU(\mathcal{A})$ such that $\Delta^* \restriction [FU(\mathcal{A}_0)]^k_< =$ const. If $\Delta^*([FU(\mathcal{A}_0)]^k_<)=\{0\}$ we are done: Δ_0 and Δ_1 are identical on $[FU(\mathcal{A}_0)]^k_<$. So let $\Delta^*([FU(\mathcal{A}_0)]^k_<)=\{1\}$. By a construction similar to that used in Case A we get an ascending collection $\mathcal{B}\subset FU(\mathcal{A}_0)$ such that $\Delta_0([FU(\mathcal{B})]^k_<)\cap\Delta_1([FU(\mathcal{B})]^k_<)=\emptyset$. We omit the details. This finishes the proof of Lemma 2.2. □

Proof of Theorem 1.6. The proof is by induction on k. For $k=1$ we have Taylor's result. Let $k\geqq 2$ and assume that the theorem is valid for all ascending collections $\mathcal{A}$ and every $k'<k$. Let $\mathcal{A}$ be an ascending collection and let $\Delta: [FU(\mathcal{A})]^k_< \to \omega$ be a mapping.

Lemma 2.3. *Let $k\geqq 2$ be an integer and let $\mathcal{A}$ be an ascending collection. Then for every mapping $\Delta: [FU(\mathcal{A})]^k_< \to\omega$ there exist an ascending collection $\mathcal{B}\subseteqq FU(\mathcal{A})$ and a k-canonical sequence $((T_i,f_i))_{i<k_0}$ with $\bigcup_{i<k_0} T_i\subseteqq \{1,\dots,k-1\}$ such that for every $S\in FU(\mathcal{B})$ and every $\{S_1,\dots,S_{k-1}\}_<$, $\{S_1^*,\dots,S_{k-1}^*\}_<\in[FU(\{B\in\mathcal{B}\mid S<B\})]^{k-1}_<$ it is valid:*

$$\Delta(\{S,S_1,\dots,S_{k-1}\}_<)=\Delta(\{S,S_1^*,\dots,S_{k-1}^*\}_<)$$

iff

$$\{f_i(\bigcup_{j\in T_i} S_j)\mid i<k_0\}=\{f_i(\bigcup_{j\in T_i} S_j^*)\mid i<k_0\}.$$

Proof of Lemma 2.3. Let $\Delta: [FU(\mathcal{A})]^k_<\to\omega$ be given. For $S\in FU(\mathcal{A})$ let $\Delta_S: [FU(\{A\in\mathcal{A}\mid S<A\})]^{k-1}_<\to\omega$ be a mapping defined by $\Delta_S(\{S_0,\dots,S_{k-2}\}_<)=\Delta(\{S,S_0,\dots,S_{k-2}\})$. Put $B_{-1}:=\min\mathcal{A}$ and $\mathcal{B}^{(-1)}:=\mathcal{A}\setminus\{\min\mathcal{A}\}$. We will construct a sequence $(B_i)_{i<\omega}$ of finite nonempty sets and a sequence $(\mathcal{B}^{(i)})_{i<\omega}$ of ascending collections with the following properties. For every $i<\omega$:

(2.3.1) $\quad B_i\in FU(\mathcal{B}^{(i-1)})$ and $\mathcal{B}^{(i)}\subset FU(\mathcal{B}^{(i-1)})$;

(2.3.2) $\quad B_{i-1}<B_i$ and $B_i<\min\mathcal{B}^{(i)}$;

(2.3.3) $\quad$ For every $S\in FU(\{B_0,\dots,B_i\})$ with $B_i\subseteqq S$ there is a $(k-1)$-canonical sequence $((T_i^S,f_i^S))_{i<k_0(S)}$ such that type $(\Delta_S\restriction[FU(\mathcal{B}^{(i)})]^{k-1}_<)=((T_i^S,f_i^S))_{i<k_0(S)}$.

Assume that for some n, finite nonempty sets B_i and ascending collections $\mathcal{B}^{(i)}$ are constructed such that (2.3.1) up to (2.3.3) are valid for $i < n$.

Step n: Put $B_n := \min \mathcal{B}^{(n-1)}$. For $S \in FU(\{B_0, \ldots, B_n\})$ with $B_n \subseteqq S$ consider the mapping $\Delta_S\restriction[FU(\mathcal{B}^{(n-1)}\setminus\{B_n\})]^{k-1}_<$. Then 2^n applications of the induction hypothesis yield an ascending collection $\mathcal{B}^{(n)} \subseteqq FU(\mathcal{B}^{(n-1)} \setminus \{B_n\})$ and for every set $S \in FU(\{B_0, \ldots, B_n\})$ with $B_n \subseteqq S$ a $(k-1)$-canonical sequence $((T_i^S, f_i^S))_{i<k_0(S)}$ such that type $(\Delta_S\restriction[FU(\mathcal{B}^{(n)})]^{k-1}_<) = ((T_i^S, f_i^S))_{i<k_0(S)}$.

Finally, we get a sequence $(B_i)_{i<\omega}$ of finite nonempty sets and a sequence $(\mathcal{B}^{(i)})_{i<\omega}$ of ascending collections such that (2.3.1) up to (2.3.3) are valid for $i < \omega$. Put $\hat{\mathcal{B}} := \{B_0, B_1, \ldots, \}_<$. Let $\mathcal{T} = \{((T_i^S, f_i^S))_{i<k_0(S)} \mid S \in FU(\hat{\mathcal{B}})\}$ be the set of arising types. Clearly $\mathcal{T}$ is a finite set. Define a mapping $\Delta'\colon FU(\hat{\mathcal{B}}) \to \mathcal{T}$ by $\Delta'(S) = ((T_i^S, f_i^S))_{i<k_0(S)}$. By Theorem 1.5 there is an ascending collection $\mathcal{B} \subseteqq FU(\hat{\mathcal{B}})$ and a $(k-1)$-canonical sequence $((T_i^*, f_i))_{i<k_0} \in \mathcal{T}$ such that $\Delta'((FU(\mathcal{B})) = \{((T_i^*, f_i))_{i<k_0}\}$. Putting $T_i := \{t+1 \mid t \in T_i^*\}$ for $i < k_0$ finishes the proof. $\square$

Let $\mathcal{B} \subseteqq FU(\mathcal{A})$ be an ascending collection and let $((T_i, f_i))_{i<k_0}$ be a k-canonical sequence according to Lemma 2.3. If $k_0 < k-1$, the mapping Δ induces another mapping $\Delta'\colon [FU(\mathcal{B})]^{k_0-1}_< \to \omega \cup \{-1\}$ by

$$\Delta'(\{S_0, \ldots, S_{k_0}\}_<) = \begin{cases} \Delta(\{S_0^*, S_1^*, \ldots, S_{k-1}^*\}_<) \text{ if there exists} \\ \qquad \{S_1^*, \ldots, S_{k-1}^*\}_< \in [FU(\mathcal{B})]^{k-1}_< \\ \qquad \text{such that } \{S_i \mid 1 \leqq i \leqq k_0\} = \{\bigcup_{j \in T_i} S_j^* \mid i < k_0\}, \\ -1 \quad \text{else.} \end{cases}$$

By choice of $\mathcal{B}$ and $((T_i, f_i))_{i<k_0}$ the mapping Δ' is well defined. The induction hypothesis applied to Δ' yields an ascending collection $\mathcal{C} \subseteqq FU(\mathcal{B})$ and a (k_0+1)-canonical sequence $((T_i', f_i'))_{i<k'}$ such that type $(\Delta'\restriction[FU(\mathcal{C})]_<{}^{k_0+1}) = ((T_i', f_i'))_{i<k'}$. Transforming the sequence $((T_i', f_i'))_{i<k'}$ according to the amalgamation via Δ', we are done.

Now let $k_0 = k-1$. For technical reasons put $f_{i+1} := f_i$ for $i < k_0$. Thus we have:

(2.3.4) (i) $((\{i\}, f_i))_{1 \leqq i < k}$ is a k-canonical sequence;

(ii) For every $\{S_0, S_1, \ldots, S_{k-1}\}_<, \{S_0, S_1^*, \ldots, S_{k-1}^*\}_< \in [FU(\mathcal{B})]^k_<$:

$$\Delta(\{S_0, S_1, \ldots, S_{k-1}\}_<) = \Delta(\{S_0, S_1^*, \ldots, S_{k-1}^*\}_<)$$

iff

$$\{f_i(S_i) \mid 1 \leqq i < k\} = \{f_i(S_i^*) \mid 1 \leqq i < k\}.$$

Let $\Delta^*\colon [FU(\mathcal{B})]^{k+1}_{<} \to \{0,1\}$ be a mapping defined by

$$\Delta^*(\{S_0,\dots,S_k\}_<)=0$$

iff

$$\Delta(\{S_0\cup S_1, S_2,\dots,S_k\}_<)=\Delta(\{S_0, S_1\cup S_2, S_3,\dots,S_k\}_<).$$

By Theorem 1.5 there exists an ascending collection $\mathcal{C}\subseteqq FU(\mathcal{B})$ such that $\Delta^*([FU(\mathcal{C})]^{k+1}_<)=\{c\}$ for some $c<2$. For $c=0$ we are done by induction hypothesis, so let $c=1$, i.e.,

$$\text{(2.3.5)}\qquad \begin{array}{l}\text{For every } \{S_0,\dots,S_k\}_<\in[FU(\mathcal{C})]^{k+1}_<:\\ \Delta(\{S_0\cup S_1, S_2,\dots,S_k\}_<)\neq\Delta(\{S_0, S_1\cup S_2, S_3,\dots,S_k\}_<).\end{array}$$

Thus entries in components 0 and 1 cannot be shifted without changing the image.

In the following we examine the pattern of Δ w.r.t. the smallest element of k-term ascending families. For that purpose consider the mappings $\Delta_F\colon[FU(\mathcal{C})]^{k+1}_<\to\{0,1\}$ and $\Delta_L\colon[FU(\mathcal{C})]^{k+1}_<\to\{0,1\}$ and $\Delta_{FL}\colon[FU(\mathcal{C})]_<{}^{k+2}\to\{0,1\}$ defined by

$$\Delta_F(\{S_0,\dots,S_k\}_<)=0$$

iff

$$\Delta(\{S_0\cup S_1,S_2,\dots,S_k\}_<)=\Delta(\{S_0,S_2,\dots,S_k\}_<),$$

$$\Delta_L(\{S_0,\dots,S_k\}_<)=0$$

iff

$$\Delta(\{S_0\cup S_1,S_2,\dots,S_k\}_<)=\Delta(\{S_1,S_2,\dots,S_k\}_<),$$

$$\Delta_{FL}(\{S_0,\dots,S_{k+1}\}_<)=0$$

iff

$$\Delta(\{S_0\cup S_1\cup S_2,S_3,\dots,S_{k+1}\}_<)=\Delta(\{S_0\cup S_2,S_3,\dots,S_{k+1}\}_<).$$

By Theorem 1.5 there are an ascending collection $\mathcal{D}\subseteqq FU(\mathcal{C})$ and numbers $c_F,c_L,c_{FL}<2$ such that

$$\Delta_F([FU(\mathcal{D})]^{k+1}_<)=\{c_F\},$$

$$\Delta_L([FU(\mathcal{D})]^{k+1}_<)=\{c_L\}$$

and

$$\Delta_{FL}([FU(\mathcal{D})]^{k+2}_<)=\{c_{FL}\}.$$

Since $c_F{=}0$ or $c_L{=}0$ implies $c_{FL}{=}0$, only five possible values for (c_F,c_L,c_{FL}) remain, namely $(0,0,0)$, $(0,1,0)$, $(1,0,0)$, $(1,1,0)$ and $(1,1,1)$. According to their values we distinguish five cases.

$(c_F, c_L, c_{FL}) = (0,0,0)$. We claim that

$$\Delta(\{S_0, S_1, \dots, S_{k-1}\}_<) = \Delta(\{S_0^*, S_1, \dots, S_{k-1}\}_<)$$

for every $\{S_0, S_1, \dots, S_{k-1}\}_<$, $\{S_0^*, S_1, \dots, S_{k-1}\}_< \in [FU(\mathcal{D})]^k_<$. This can be seen as follows: Let $D_0 = \min \mathcal{D}$, then

$$\begin{aligned}
&\Delta(\{S_0, S_1, \dots, S_{k-1}\}_<) \\
&= \Delta(\{D_0 \cup S_0, S_1, \dots, S_{k-1}\}_<) && \text{since } c_L = 0, \\
&= \Delta(\{D_0, S_1, \dots, S_{k-1}\}_<) && \text{since } c_F = 0, \\
&= \Delta(\{D_0 \cup S_0^*, S_1, \dots, S_{k-1}\}_<) && \text{since } c_F = 0. \\
&= \Delta(\{S_0^*, S_1, \dots, S_{k-1}\}_<) && \text{since } c_L = 0.
\end{aligned}$$

By (2.3.4) for all $\{S_0, \dots, S_{k-1}\}_<$, $\{S_0^*, \dots, S_{k-1}^*\}_< \in [FU(\mathcal{D})]^k_<$, where w.l.o.g. $\max S_0 \leqq \max S_0^*$, it is valid:

$$\Delta(\{S_0, S_1, \dots, S_{k-1}\}_<) = \Delta(\{S_0^*, S_1^*, \dots, S_{k-1}^*\}_<)$$

iff

$$\Delta(\{S_0, S_1, \dots, S_{k-1}\}_<) = (\{S_0, S_1^*, \dots, S_{k-1}^*\}_<)$$

iff

$$\{f_i(S_i) \mid 1 \leqq i \leqq k-1\} = \{f_i(S_i^*) \mid 1 \leqq i \leqq k-1\}.$$

We use the following propositions for the remaining four cases:

PROPOSITION 2.4. *Let $c_L = 1$. Then there exists an ascending collection $\mathcal{E} \subset FU(\mathcal{D})$ such that for every $\{S_0, \dots, S_{k-1}\}_<$, $\{S_0^*, \dots, S_{k-1}^*\}_< \in [FU(\mathcal{E})]^k_<$:*

$$\Delta(\{S_0, \dots, S_{k-1}\}_<) = \Delta(\{S_0^*, \dots, S_{k-1}^*\}_<)$$

implies

$$\min S_0 = \min S_0^*.$$

PROPOSITION 2.5. *Let $c_F = 1$. Then there exists an ascending collection $\mathcal{E} \subset FU(\mathcal{D})$ such that for every $\{S_0, \dots, S_{k-1}\}_<$, $\{S_0^*, \dots, S_{k-1}^*\} \in [FU(\mathcal{E})]^k_<$:*

$$\Delta(\{S_0, \dots, S_{k-1}\}_<) = \Delta(\{S_0^*, \dots, S_{k-1}^*\}_<)$$

implies

$$\max S_0 = \max S_0^*.$$

PROPOSITION 2.6. *Let $c_{FL} = 1$. Then there exists an ascending collection $\mathcal{E} \subset FU(\mathcal{D})$ such that for every $\{S_0, \dots, S_{k-1}\}_<$, $\{S_0^*, \dots, S_{k-1}^*\}_< \in [FU(\mathcal{E})]^k_<$ with $\max S_0 = \max S_0^*$:*

$$\Delta(\{S_0, \dots, S_{k-1}\}_<) = \Delta(\{S_0^*, \dots, S_{k-1}^*\}_<)$$

implies

$$S_0 = S_0^*.$$

The proofs of these propositions are quite technical, however, they are along the lines of the proof of Lemma 2.2, so we omit them here. For details compare [3].

$(c_F, c_L, c_{FL}) = (0,1,0)$. Let $\mathcal{E}$ be an ascending collection according to Proposition 2.4. Since $c_F = 0$ and (2.3.4) holds, we have for every $\{S_0, \ldots, S_{k-1}\}_<$, $\{S_0^*, \ldots, S_{k-1}^*\}_< \in [FU(\mathcal{E})]^k_<$:

$$\Delta(\{S_0, \ldots, S_{k-1}\}_<) = \Delta(\{S_0^*, \ldots, S_{k-1}^*\}_<)$$

iff

$$\min S_0 = \min S_0^*$$

and

$$\{f_i(S_i) \mid 1 \leqq i \leqq k-1\} = \{f_i(S_i^*) \mid 1 \leqq i \leqq k-1\}.$$

$(c_F, c_L, c_{FL}) = (1,0,0)$. Let $\mathcal{E}$ be an ascending collection according to Proposition 2.5. Since $c_L = 0$ and (2.3.4) is valid, we have for every $\{S_0, \ldots, S_{k-1}\}_<$, $\{S_0^*, \ldots, S_{k-1}^*\}_< \in [FU(\mathcal{E})]^k_<$:

$$\Delta(\{S_0, \ldots, S_{k-1}\}_<) = \Delta(\{S_0^*, \ldots, S_{k-1}^*\}_<)$$

iff

$$\max S_0 = \max S_0^*$$

and

$$\{f_i(S_i) \mid 1 \leqq i \leqq k-1\} = \{f_i(S_i^*) \mid 1 \leqq i \leqq k-1\}.$$

$(c_F, c_L, c_{FL}) = (1,1,0)$. Let $\mathcal{E}$ be an ascending collection according to Propositions 2.4 and 2.5. Since $c_{FL} = 0$ and (2.3.4) is valid, we have for every $\{S_0, \ldots, S_{k-1}\}_<$, $\{S_0^*, \ldots, S_{k-1}^*\}_< \in [FU(\mathcal{E})]^k_<$:

$$\Delta(\{S_0, \ldots, S_{k-1}\}_<) = \Delta(\{S_0^*, \ldots, S_{k-1}^*\}_<)$$

iff

$$\min S_0 = \min S_0^* \quad \text{and} \quad \max S_0 = \max S_0^*$$

and

$$\{f_i(S_i) \mid 1 \leqq i \leqq k-1\} = \{f_i(S_i^*) \mid 1 \leqq i \leqq k-1\}.$$

$(c_F, c_L, c_{FL}) = (1,1,1)$. Let $\mathcal{E}$ be an ascending collection according to Propositions 2.5 and 2.6. By (2.3.4) for every $\{S_0, ..., S_{k-1}\}_<$, $\{S_0^*, ..., S_{k-1}^*\}_<$ $\in [FU(\mathcal{E})]^k_<$ it is valid

$$\Delta(\{S_0, \ldots, S_{k-1}\}_<) = \Delta(\{S_0^*, \ldots, S_{k-1}^*\}_<)$$

iff

$$S_0 = S_0^* \text{ and } \{f_i(S_i) \mid 1 \leqq i \leqq k-1\} = \{f_i(S_i^*) \mid 1 \leqq i \leqq k-1\}.$$

This finishes the proof of Theorem 1.6. □

The proof of Theorem 1.7 is a straightforward application of Theorem 1.6 and, eventually after merging entries of coordinates, of Lemma 2.2.

REFERENCES

[1] Erdős, P. and Rado, R., A combinatorial theorem, *J. London Math. Soc.* **25** (1950), 249–255. *MR* **12**-322

[2] Hindman, N., Finite sums from sequences within cells of a partition of N, *J. Combinatorial Theory Ser. A* **17** (1974), 1–11. *MR* **50** #2067

[3] Lefmann, H., Kanonische Partitionssätze, Dissertation, Bielefeld, 1985.

[4] Milliken, K. R., Ramsey's theorem with sums or unions, *J. Combinatorial Theory Ser. A* **18** (1975), 276–290. *MR* **51** #10106

[5] Prömel, H. J. and Voigt, B., Canonizing Ramsey theorems for finite graphs and hypergraphs, *Discrete Math.* **54** (1985), 49–59. *MR* **87d**:05121

[6] Ramsey, F. P., On a problem of formal logic, *Proc. London Math. Soc.* (2) **30** (1930), 264–286. *Jb. Fortschritte Math.* **55**, 32–33

[7] Taylor, A. D., A canonical partition relation for finite subsets of ω, *J. Combinatorial Theory Ser. A* **21** (1976), 137–146. *MR* **54** #12530

[8] Voigt, B., Canonizing partition theorems: diversification, products, and iterated versions, *J. Combinatorial Theory Ser. A* **40** (1985), 349–376. *MR* **87h**:05028

(Received December 18, 1988)

UNIVERSITÄT DORTMUND
FACHBEREICH INFORMATIK
LEHRSTUHL II
D–44221 DORTMUND
FEDERAL REPUBLIC OF GERMANY

lefmann@ls2.informatik.uni-dortmund.de

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 375–384

WEIGHTED NIKOLSKIĬ-TYPE INEQUALITIES FOR TRIGONOMETRIC POLYNOMIALS

N. X. KY

1. Introduction

Let $L^p = L^p[0, 2\pi]$ $(1 \leqq p \leqq \infty)$ be the space of functions defined on $[0, 2\pi]$ with the norm

$$\|f\|_p = \left\{ \int_0^{2\pi} |f(x)|^p \right\}^{\frac{1}{p}} \qquad (1 \leqq p < \infty),$$

$$\|f\|_\infty = \operatorname*{ess.sup}_{x \in [0,2\pi]} |f(x)|.$$

Denote by T_n the set of all trigonometric polynomials of order at most n $(n = 1, 2 \ldots)$. Nikolskiĭ [12] proved the inequality

$$\|t_n\|_q \leqq c(p,q) n^{\frac{1}{p} - \frac{1}{q}} \|t_n\|_p \tag{1}$$

$(1 \leqq p < q \leqq \infty,\ t_n \in T_n,\ n = 1, 2, \ldots)$, where $c(p,q)$ is a constant depending only on p and q. The inequality is sharp in the sense that the constant $c(p,q)$ cannot be replaced by $o(1) = o_n(1)$ $(n \to \infty)$. There are different proofs and generalizations for the above inequality (see e.g. [1], [10], etc.).

Recently, inequalities of Nikolskiĭ-type for other systems of functions were considered by many authors, for example [4], [5], [11], [13].

Let t_n^* be the decreasing rearrangement function of $t_n \in T_n$ (for the definition, see [2, p. 7]). Then (1) can be written in the form

$$\|t_n^*\|_q \leqq c(p,q) n^{\frac{1}{p} - \frac{1}{q}} \|t_n\|_p. \tag{2}$$

In Part 2 of this note, we shall consider inequalities of types (1) and (2) in the cases, when t_n and t_n^* on the left-hand sides are replaced by $t_n u$ and $t_n^* u$, respectively, with the weight function u satisfying some conditions. In Part 3, analogous inequalities are given in Lorentz spaces. The method considered here can be applied to the cases of algebraic polynomials on finite and infinite intervals, which will be detailed in a forthcoming paper.

1991 *Mathematics Subject Classification.* Primary 41A17.
Key words and phrases. Weighted Nikolskiĭ-type inequalities.

Research supported by the Hungarian National Foundation for Scientific Research Grant No. 1801.

0081–6906/96/$ 5.00 ©1996 *Akadémiai Kiadó, Budapest*

2. Nikolskiĭ-type inequalities with weights

Let $1 \leqq \lambda,\ q < \infty,\ \frac{1}{q} + \frac{1}{q'} = 1$. We introduce the following transforms of functions:

$$F_{q,\varepsilon}[u,t] := \left(\int_0^t u_\varepsilon^q(x)dx \right) u_\varepsilon^{-\frac{q}{q'}}(t) \tag{3}$$

($u_\varepsilon(x) := u(x) + \varepsilon$, $u(x) \geqq 0$, $x \in [0, 2\pi]$, $u \in L^q$ and $\varepsilon > 0$) and

$$F^{\langle\lambda\rangle}(t) := \frac{F(t)}{t^{\frac{1}{\lambda}}} \tag{4}$$

for any function F defined on $[0, 2\pi]$.

It is clear that for any $u(x) \geqq 0$, $u \in L^q$ and $\varepsilon > 0$, we have $F_{q,\varepsilon}[u,t] \geqq 0$ and $F_{q,\varepsilon}[u,t]$ is bounded on $[0,2\pi]$, whence $F_{q,\varepsilon}[u,t] \in L^q$.

Let $B_{q,\lambda}$ be the collection of all functions $u \in L^q$ ($u(x) \geqq 0$) for which $F_{q,\varepsilon}^{\langle\lambda\rangle}[u,t] \in L^q$ ($\forall \varepsilon > 0$). Then by the above arguments we have that if $u \in$ $\in B_{q,\lambda}$ then $F_{q,\delta}\left|F_{q,\varepsilon}^{\langle\lambda\rangle}[u],t\right| \in L^q$ for every $\varepsilon, \delta > 0$.

Now, for $n = 1, 2, \dots$ and $u \in B_{q,\lambda}$, define

$$\varrho_n(u,\lambda,q) := \inf_{\substack{\varepsilon>0 \\ \delta>0}} \left\{ \int_{\frac{1}{n}}^{2\pi} \left(\frac{F_{q,\delta}[F_{q,\varepsilon}^{\langle\lambda\rangle}[u],t]}{t^2} \right)^q dt \right\}^{\frac{1}{q}}. \tag{5}$$

For a function $f(x)$ measurable on $[0, 2\pi]$, let f^* be the decreasing rearrangement function of f.

THEOREM 1. *Let $1 \leqq \lambda,\ q < \infty$ and $t_n \in T_n$ $(n \geqq 1)$. Then:*

1. *For any $u \in B_{q,\lambda}$ we have*

$$||t_n^* u||_q \leqq c(u,\lambda,q)[\varrho_n(u,\lambda,q) + 1]||t_n||_\lambda; \tag{6}$$

2. *If $u^* \in B_{q,\lambda}$ then*

$$||t_n u||_q \leqq c(u,\lambda,q)[\varrho_n(u^*,\lambda,q) + 1]||t_n||_\lambda, \tag{7}$$

where $\varrho_n(u,\lambda,q)$ is defined by (5). Here and later $c(.,.,\dots)$ denotes a constant depending only on the variables specified in the brackets.

REMARK 1. a. Set $u(t) \equiv 1$. By simple computation we get

$$\varrho_n(u,\lambda,q) \sim n^{\frac{1}{\lambda} - \frac{1}{q}} \qquad (1 \leqq \lambda < q < \infty).$$

This fact shows that inequality (7) is a generalization of the Nikolskiĭ inequality in the case $1 \leqq \lambda < q < \infty$.

b. One may ask whether inequalities (6) and (7) are sharp in the sense stated above. In the case of general weights u, this problem is still open. In the following we shall show an example, different from Case a, when inequality (7) is sharp.

Let $u(t) = t^{-\alpha}$, $0 \leqq \alpha q < 1 \leqq q < \infty$. It is easy to compute that

$$\varrho_n(u, \lambda, q) = \varrho_n(u^*, \lambda, q) \sim n^{\frac{1}{\lambda} - \frac{1}{q} + \alpha}$$
$$(1 \leqq \lambda, q < \infty;\ \frac{1}{\lambda} - \frac{1}{q} + \alpha > 0).$$

Therefore using Theorem 1 we have the inequality

$$\|t_n(x)x^{-\alpha}\|_q \leqq c(\alpha, \lambda, q) n^{\frac{1}{\lambda} - \frac{1}{q} + \alpha} \|t_n\|_\lambda \tag{8}$$
$$\left(1 \leqq \lambda,\ q < \infty;\ 0 \leqq \alpha q < 1,\ \frac{1}{\lambda} - \frac{1}{q} + \alpha > 0,\ t_n \in T_n\right).$$

Now, let us prove that inequality (8) is indeed sharp. For this purpose we can choose a test polynomial, namely the Fejér kernel

$$K_n(t) = \left[\frac{\sin \frac{n+1}{2} t}{\sin \frac{t}{2}}\right]^2.$$

This is an even trigonometric polynomial of order n. We have

$$\|K_n(t)t^{-\alpha}\|_q^q = 2\int_0^\pi K_n^q(t) t^{-\alpha q}\, dt =$$
$$= 2\left[\int_0^{\frac{1}{n}} K_n^q(t) t^{-\alpha q}\, dt + \int_{\frac{1}{n}}^{\pi} K_n^q(t) t^{-\alpha q}\, dt\right] =$$
$$=: I_1(n) + I_2(n).$$

It is easy to see that

$$I_1(n) \sim n^{2q} \int_0^{\frac{1}{n}} t^{-\alpha q}\, dt \sim n^{\alpha q + 2q - 1},$$

$$I_2(n) \leqq c(\alpha, q)\left[\int_{\frac{1}{n}}^{\pi} \frac{1}{t^{2q+\alpha q}}\, dt\right] \sim n^{\alpha q + 2q - 1}.$$

Consequently,

$$\|K_n(t)t^{-\alpha}\|_q \sim n^{\alpha+2-\frac{1}{q}}$$

and so we have for $1 \leqq \lambda,\ q < \infty;\ \frac{1}{\lambda} - \frac{1}{q} + \alpha > 0$

$$\|K_n(t)t^{-\alpha}\|_q \sim n^{\frac{1}{\lambda}-\frac{1}{q}+\alpha}\|K_n\|_\lambda,$$

which proves that inequality (8) is sharp.

For the proof of Theorem 1 we need some lemmas.

LEMMA 1. *Let* $1 \leqq q < \infty,\ \frac{1}{q} + \frac{1}{q'} = 1$. *Let* $u(x) \geqq 0,\ u \in L^q$ *and* $\varepsilon > 0$. *Introduce*

$$v_{(\varepsilon)}(t) := \left(\frac{q'}{q}\right)^{-\frac{1}{q'}} F_{q,\varepsilon}[u,t] \tag{9}$$

where $F_{q,\varepsilon}$ *is defined by* (3) *and* $(q'/q)^{-\frac{1}{q'}} := 1$ *if* $q = 1$. *Then*

$$\sup_{0<r<2\pi} \left(\int_0^r u_\varepsilon^q(x)dx\right)^{\frac{1}{q}} \left(\int_r^{2\pi} v_{(\varepsilon)}^{-q'}(x)dx\right)^{\frac{1}{q'}} \leqq 1. \tag{10}$$

PROOF. Case A. $1 < q < \infty$. We introduce the function

$$E(t) := \left\{\int_0^t u_\varepsilon^q(x)dx\right\}^{-\frac{q'}{q}}. \tag{11}$$

Then by a well-known property of the integral functions we have

$$v_{(\varepsilon)}^{-q'}(t) = \frac{q'}{q}\left(\int_0^t u_\varepsilon^q(x)dx\right)^{-\frac{q'}{q}-1} u_\varepsilon^q(t) = -E'(t) \text{ a.e. on } [0,2\pi]. \tag{12}$$

On the other hand, since the function

$$\left(\int_0^t u_\varepsilon^q(x)dx\right)^{-\frac{q'}{q}-1}$$

is decreasing on $(0,2\pi)$, it is bounded on any interval $[r,2\pi]$ $(0<r<2\pi)$, and so $v_{(\varepsilon)}^{-q'}$ is integrable over the interval $[r,2\pi]$. However, $E(t)$ is absolutely continuous on $[r,2\pi]$, therefore we have by (12)

$$\int_r^{2\pi} v_{(\varepsilon)}^{-q'}(x)dx = E(r) - E(2\pi) \leqq E(r) = \left(\int_0^r u_\varepsilon^q(x)dx\right)^{-\frac{q'}{q}} \quad (0<r<2\pi),$$

and so

$$\left(\int_u^r u_\varepsilon^q(x)dx\right)^{\frac{1}{q}}\left(\int_r^{2\pi} v_{(\varepsilon)}^{-q'}(x)dx\right)^{\frac{1}{q'}} \leqq 1 \quad (0<r<2\pi),$$

which proves (10).

Case B. $q=1$. Then by (3) and (9) we have

$$v_{(\varepsilon)}(t)=\int_0^t u_\varepsilon^q(x)dx.$$

Consequently, for $0<r<2\pi$

$$\sup_{t\in[r,2\pi]}\frac{1}{v_{(\varepsilon)}(t)}=\frac{1}{\int_0^r u_\varepsilon^q(x)dx},$$

which implies the case $q=1$ of (10). The proof of Lemma 1 is complete.

In the following we shall use the following notions for any function $f\in L^\lambda$

$$E_k(f)_\lambda := \inf_{t_k\in T_k}\|f-t_k\|_\lambda \qquad (k=0,1,\dots)$$

$$\omega(f,\delta)_\lambda := \sup_{0<h\leqq\delta}\left\{\int_0^{2\pi-h}|f(x+h)-f(x)|^\lambda dx\right\}^{\frac{1}{\lambda}} \qquad (0<\delta<2\pi).$$

Lemma 2. *Let $f\in L^\lambda$ $(1\leqq\lambda<\infty)$. Define*

$$\varphi(t):=\varphi_{f,\lambda}(t):=\begin{cases} E_{k-1}(f)_\lambda & \text{for } t\in\left[\frac{1}{k+1},\frac{1}{k}\right]\\ \varphi(1) & \text{for } 1\leqq t\leqq 2\pi.\end{cases}$$

Then

$$\omega(f,t)_\lambda \leqq c(x)t\int_t^{2\pi}\frac{\varphi(x)}{x^2}dx \qquad (0<t<2\pi). \tag{13}$$

Proof. (13) indeed follows from the well-known converse theorem concerning best approximation by trigonometric polynomials in L^λ-spaces (see e.g. [14, p. 344, Theorem 6.1.1]).

Lemma 3. *Let $f\in L^\lambda$ $(1\leqq\lambda<\infty)$. Then*

$$f^*(x)\leqq c(\lambda)\left\{\int_x^{2\pi}\frac{\omega(f,t)_\lambda}{t^{1+\frac{1}{\lambda}}}dt+\|f\|_\lambda\right\} \qquad (0<x<2\pi). \tag{14}$$

PROOF. This inequality follows from [6, Corollary 3.4] and [7, Theorem A].

PROOF of Theorem 1.1. Inequality (6): Let $u \in B_{q,\lambda}$, $t_n \in T_n$, $\varepsilon>0$. Since the functions u_ε and $v_{(\varepsilon)}$ defined by (9) satisfy (10), using [8, Theorem 2], we have by (13) and (14)

$$\|t_n^* u\|_q \leqq \|t_n^* u_\varepsilon\|_q \leqq$$

$$\leqq c(\lambda)\left[\left\|u_\varepsilon(t)\int_t^{2\pi}\frac{\omega(t_n,x)_\lambda}{x^{1+\frac{1}{\lambda}}}dt\right\|_q + \|u_\varepsilon\|_q\|t_n\|_\lambda\right] \leqq$$

$$\leqq c(\lambda,q)\left[\left\|v_{(\varepsilon)}(t)\frac{\omega(t_n,t)_\lambda}{t^{1+\frac{1}{\lambda}}}\right\|_q + \|u_\varepsilon\|_q\|t_n\|_\lambda\right] \leqq$$

$$\leqq c(\lambda,q)\left[\left\|\frac{F_{q,\varepsilon}[u,t]}{t^{1+\frac{1}{\lambda}}}t\int_t^{2\pi}\frac{\varphi_{t_n,\lambda}(x)}{x^2}dx\right\|_q + \|u_\varepsilon\|_q\|t_n\|_\lambda\right] =$$

$$= c(\lambda,q)\left[\left\|F_{q,\varepsilon}^{\langle\lambda\rangle}[u,t]\int_t^{2\pi}\frac{\varphi_{t_n,\lambda}(x)}{x^2}dx\right\|_q + \|u_\varepsilon\|_q\|t_n\|_\lambda\right].$$

Again, using [8, Theorem 2] for the weight $F_{q,\varepsilon}^{\langle\lambda\rangle}[u,t]$ we have for any $\delta>0$

$$\|t_n^* u\|_q \leqq c(\lambda,q)\left[\left\|\frac{F_{q,\delta}[F_{q,\varepsilon}^{\langle\lambda\rangle}[u,t]]}{t^2}\varphi_{t_n,\lambda}(t)\right\|_q + \|u_\varepsilon\|_q\|t_n\|_\lambda\right]. \tag{15}$$

However, by definition,

$$\varphi_{t_n,\lambda}(t) = \begin{cases} 0 & \text{for } t \leqq \frac{1}{n} \\ \leqq \|t_n\|_\lambda & \text{for } 0 < t \leqq 2\pi. \end{cases}$$

Therefore by (15)

$$\|t_n^* u\|_q \leqq c(\lambda,q)\left[\left\{\int_{\frac{1}{n}}^{2\pi}\left(\frac{F_{q,\delta}[F_{q,\varepsilon}[u,t]]}{t^2}\right)^q dt\right\}^{\frac{1}{q}} + \|u_\varepsilon\|_q\right]\|t_n\|_\lambda$$

from which (6) follows, since ε and δ are arbitrary positive numbers.

2. Inequality (7): By an inequality of Hardy and Littlewood (see [4], p. 102) we have for any $t_n \in T_n$ and $u \in L^q$ $(1 \leqq q < \infty)$

$$\|t_n u\|_q \leqq \|t_n^* u^*\|_q,$$

therefore we get (7) by taking u^* instead of u in (6). This completes the proof of the theorem.

3. Weighted Nikolskiĭ-type inequality in Lorentz spaces

Using some results of the theory of interpolation spaces we can get Nikolskiĭ-type inequalities between some Banach spaces. Some results of this type can be found in [10]. In this part we shall state a Nikolskiĭ-type inequality in Lorentz spaces with a concrete weight, and notice that the method we are applying here can be used for general weights.

Let $L_{p,q}$ be the Lorentz space of functions defined on $[0,2\pi]$ with the norm denoted by $\| \|_{p,q}$ (see e.g. [1]).

THEOREM 2. *Let $1<p_0,\ p_1<\infty,\ 0<q_0\leqq q_1\leqq\infty$ and let $\alpha\geqq 0$ be a real number such that $\alpha p_1<1$. Then we have for every $t_n\in T_n$ $(n\geqq 1)$, $\frac{1}{p_0}-\frac{1}{p_1}+\alpha>0$*

$$\|t_n(t)t^{-\alpha}\|_{p_1,q_1}\leqq c(p_0,q_0,p_1,q_1,\alpha)n^{\frac{1}{p_0}-\frac{1}{p_1}+\alpha}\|t_n\|_{p_0,q_0}. \tag{16}$$

The inequality is sharp, i.e. the constant $c(p_0,q_0,p_1,q_1,\alpha)$ cannot be replaced by $o(1)$ $(n\to\infty)$.

REMARK 2. In the case $\alpha=0$ (that is, in the case without weight), inequality (16) was proved in [10], Corollary 1 for $1<p_0<p_1<\infty$, $1\leqq q_0$, $q_1\leqq\infty$ or $p_0=1$, $q_0=\infty$ or $p_1=q_1=\infty$ without proving sharpness.

PROOF of Theorem 2. Let $\mathsf{V}_n(f)$ be the n-th de la Vallée–Poussin mean of $f\in L^1$. It is known that

$$\|\mathsf{V}_n(f)\|_p\leqq c(p)\|f\|_p \qquad (f\in L^p,\ 1\leqq p\leqq\infty,\ n=1,2\dots) \tag{17}$$

and

$$\mathsf{V}_n(t_n)=t_n \qquad \text{for every } t_n\in T_n\ (n=1,2,\dots). \tag{18}$$

Consequently, from (15) we have

$$\|\mathsf{V}_n(f,t)t^{-\alpha}\|_p\leqq c(p,\lambda,\alpha)n^{\frac{1}{\lambda}-\frac{1}{p}+\alpha}\|f\|_\lambda \tag{19}$$
$$(1\leqq\lambda,\ p<\infty,\ \alpha p<1,\ f\in L^\lambda,\ n=1,2,\dots).$$

Since the interpolation spaces between two spaces L^r and L^s are Lorentz spaces, more precisely

$$(L^r,L^s)_{\theta,q}=L_{p,q}$$
$$\left(1\leqq r\neq s<\infty,\ 0<q\leqq\infty,\ \frac{1}{p}=\frac{1-\theta}{r}+\frac{\theta}{s},\ 0<\theta<1\right)$$

(see [2, T. 5.3.1]), by (19) using [2], Theorem 3.1.2 we get

$$\|\mathsf{V}_n(f,t)t^{-\alpha}\|_{p_1,q_0}c(p_0,p_1,q_0,\alpha)n^{\frac{1}{p_0}-\frac{1}{p_1}+\alpha}\|f\|_{p_0,q_0} \tag{20}$$
$$(1<p_0,\ p_1<\infty,\ 0<q_0\leqq\infty,\ f\in L_{p_0,q_0},n=1,2,\dots).$$

On the other hand, since

$$L_{p_1,q_0} \subset L_{p_1,q_1} \qquad (q_0 < q_1)$$

(see [1]), (20) implies that

$$\begin{gathered} \|\mathsf{V}_n(f,t)t^{-\alpha}\|_{p_1,q_1} \leqq c(p_0,q_0,p_1,q_1,\alpha)n^{\frac{1}{p_0}-\frac{1}{p_1}+\alpha}\|f\|_{p_0,q_0} \\ (1<p_0,\ p_1<\infty,\ 0<q_0\leqq q_1\leqq\infty,\ f\in L_{p_0,q_0},\ n=1,2,\dots). \end{gathered} \tag{21}$$

From (18) and (21) we get (16).

It remains to prove the sharpness of the inequality (16). Recall that in the case $p_0 = q_0$, $p_1 = q_1$ this statement was proved above. We shall see that the general cases can be reduced to these ones.

Suppose first that for some $1 < p_0,\ p_1 < \infty,\ p_0 = q_0 \leqq q_1 \leqq \infty,\ p_1 \neq q_1$, inequality (16) is not sharp, which means that for every $t_n \in T_n$

$$\|\mathsf{V}_n(f,t)t^{-\alpha}\|_{p_1,q_1} \leqq o(1)n^{\frac{1}{p_0}-\frac{1}{p_1}+\alpha}\|t_n\|p_0.$$

Then by (17) we have

$$\begin{gathered} \|\mathsf{V}_n(f,t)t^{-\alpha}\|_{p_1,q_1} \leqq o(1)n^{\frac{1}{p_0}-\frac{1}{p_1}+\alpha}\|f\|_p \\ (f\in L^{p_0},\ n=1,2,\dots). \end{gathered}$$

Let $1 < p_1' < \infty;\ 0 < q_1' \leqq \infty$ be numbers satisfying $\alpha p_1' < 1$ and

$$\operatorname{sign}\left(\frac{1}{p_1'}-\frac{1}{q_1'}\right) = \operatorname{sign}\left(\frac{1}{p_1}-\frac{1}{q_1}\right).$$

Let

$$\frac{1}{\theta} := 1 + \frac{\dfrac{1}{p_1'}-\dfrac{1}{q_1'}}{\dfrac{1}{p_1}-\dfrac{1}{q_1}},$$

and

$$\left\{\begin{aligned} \frac{1}{\theta} &:= (1-\theta)\frac{1}{p_1}+\theta\frac{1}{q_1} \\ \frac{1}{q_\theta} &:= (1-\theta)\frac{1}{p_1'}+\theta\frac{1}{q_1'}. \end{aligned}\right. \tag{23}$$

Then $p_\theta = q_\theta$.

Now, by (16) we have

$$\begin{gathered} \|\mathsf{V}_n(f,t)t^{-\alpha}\|_{p_1',q_1'} \leqq c(p_0,p_1',q_1',\alpha)n^{\frac{1}{p_0}-\frac{1}{p_1'}+\alpha}\|f\|_{p_0} \\ (f\in L^p,\ n=1,2,\dots). \end{gathered}$$

Therefore again using the interpolation theorem we get by

$$\|V_n(f,t)t^{-\alpha}\|_{p_\theta} \leqq o(1) n^{\frac{1}{p_0}-\frac{1}{p_\theta}+\alpha} \|f\|_{p_0}$$

and so, using (18) we get

$$\|t_n(t)t^{-\alpha}\|_{p_\theta} \leqq o(1) n^{\frac{1}{p_0}-\frac{1}{p_\theta}+\alpha} \|t_n\|_{p_0} \qquad (t_n \in T_n,\ n = 1, 2, \ldots),$$

which contradicts the fact that (15) is sharp.

By similar interpolation arguments, using the just proved cases, we can prove the sharpness of (16) in all cases $1 \leqq p_0,\ p_1 < \infty,\ 0 < q_0 \leqq q_1 \leqq \infty$.

ACKNOWLEDGEMENT. The author would like to thank Prof. J. Szabados for his pointing out properties of the Fejér kernel used to prove the sharpness of (8).

REFERENCES

[1] BERGH, J. and LÖFSTRÖM, J., *Interpolation spaces. An introduction*, Grundlehren der mathematischen Wissenschaften, Vol. 223, Springer-Verlag, Berlin–New York, 1976. *MR* **58** #2349

[2] HUNT, R. A., On $L(p,q)$ spaces, *Enseignement Math.* (2) **12** (1966), 249–276. *MR* **36** #6921

[3] LUXEMBURG, W. A. J., Rearrangement-invariant Banach Function Spaces, Proceedings of the Symposium in Analysis, Queen's University, June 1976, 83–144.

[4] MARKETT, C., Nikolskiĭ-type inequalities for Laguerre and Hermite expansions, *Functions, series, operators*, Vol. I, II (Budapest, 1980), Colloq. Math. Soc. J. Bolyai, **35**, North-Holland, Amsterdam–New York, 1983, 811–834. *MR* **85h**:42031

[5] MHASKAR, H. N., Weighted polynomial approximation, *J. Approx. Theory* **46** (1986), 100–110. *MR* **87i**:41007

[6] MILMAN, M., Embedding of rearrangement invariant spaces in Lorentz spaces, *Acta Math. Acad. Sci. Hungar.* **30** (1977), 253–258. *MR* **57** #3837

[7] MILMAN, M., An inequality for generalized moduli of continuity, Notas de matematica 1977, N–9, 1–7.

[8] MUCKENHOUPT, B., Hardy's inequality with weights, *Studia Math.* **44** (1972), 31–38. *MR* **47** #418

[9] MUSIELAK, H., On some Nikolskiĭ and Oswald-type inequalities, *J. Approx. Theory* **50** (1987), 359–365. *MR* **88m**:26022

[10] NESSEL, R. J. and WILMES, G., Nikol'skiĭ-type inequalities in connection with regular spectral measures, *Acta Math. Acad. Sci. Hungar.* **33** (1979), 169–182. *MR* **80f**:41009

[11] NEVAI, P. and TOTIK, V., Sharp Nikolskiĭ inequalities with exponential weights, *Anal. Math.* **13** (1987), 261–267. *MR* **89h**:41034

[12] NIKOLSKIĬ, S. M., Inequalities for entire functions of finite degree and their application to the theory of differentiable functions of several variables, *Amer. Math. Soc. Transl. Ser.* 2, **80** (1969), 1–38.

[13] TIMAN, A. F., *Theory of approximation of functions of a real variable*, Gosudarstv. Izdat. Fiz.-Mat. Lit., Moscow, 1960 (in Russian). *MR* **22** #8257

(Received January 30, 1991)

INSTITUTE OF COMPUTER SCIENCE
NGHIA DO
TU LIEM
HANOI
VIETNAM

Present address:

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 385–404

A BITOPOLOGICAL VIEW OF QUASI-UNIFORM COMPLETENESS. III*

J. DEÁK

Abstract

We introduce some more bitopological notions of quasi-uniform completeness. Two of them (SF-complete, U-complete) have good properties. The corresponding notions in quasi-proximity spaces will also be discussed.

§ 10 SA-completeness

10.1 DEFINITION. A quasi-uniformity is SA-*complete* if each stable Cauchy filter pair has a cluster point. □

S-completeness clearly implies SA-completeness, which in turn implies L-completeness (since linked Cauchy filter pairs are stable, and it was enough to assume the existence of a cluster point in the definition of L-completeness). S- and SA-completeness both satisfy 1.1 a) to e). (Concerning b), d) and e): both notions lie between C- and L-completeness, thus the observation made in 4.2 can be applied.) We are going to show that there exists a basic SA-completion ${}^{\mathrm{SA}}\mathcal{U}$; 1.2 f) and g) are satisfied, but not h) and i).

10.2 DEFINITION. A filter pair $\mathfrak{f}^0$ is *free* if both $\mathfrak{f}^{-1}$ and $\mathfrak{f}^1$ are free (i.e. $\cap\mathfrak{f}^\imath = \emptyset$). □

Let now stable trace filter pairs be prescribed in $(X,\mathcal{U})$. For a filter $\mathfrak{f}$ in ${}^1X$, define

$$\mathfrak{f}^* = \{U^*[S]\colon S \in \mathfrak{f}, U \in \mathcal{U}\}, \quad U^*[S] = {}^1U[S] \cap {}^1U^{-1}[S] \cap X.$$

1991 *Mathematics Subject Classification.* Primary 54E15, 54D35; Secondary 54E05, 54E55, 54G20, 54D30.

Key words and phrases. Quasi-uniformity, quasi-proximity, bitopology, extension, completion, SA-complete, SF-complete, U-complete, R-complete, C-complete, F-complete, compact, fully free filter (pair), open/round/Cauchy/compressed/stable/linked filter pair.

Research supported by the Hungarian National Foundation for Scientific Research Grant No. 2114.

* See Part I (Studia Sci. Math. Hungar. 30 (1995), 389–409) for §§0 ... 4 and the references, and Part II (Studia Sci. Math. Hungar. 30 (1995), 411–431) for §§4 ... 9.

0081–6906/96/\$ 5.00 ©1996 *Akadémiai Kiadó, Budapest*

LEMMA. *Let $\mathfrak{f}$ be a free stable filter in ${}^1X$. Then*
a) $\mathfrak{f}^*$ *is a filter base;*
b) $\mathfrak{f}$ *and* $\mathfrak{f}^*$ *have the same* ${}^1\mathcal{U}$*-envelope;*
c) *if* $\mathfrak{f}^*$ ${}^1\mathcal{U}^{\mathrm{tp}}$*-converges or* ${}^1\mathcal{U}^{\mathrm{tp}}$*-clusters to some point then so does* $\mathfrak{f}$.

PROOF. a) If $T \subset S$ and $V \subset U$ then $V^*[T] \subset U^*[S]$; thus it is enough to show that $\emptyset \notin \mathfrak{f}^*$. Given $S \in \mathfrak{f}$ and $U \in \mathcal{U}$, take $V \in \mathcal{U}$ with ${}^1V^2 \subset {}^1U$. As $\mathfrak{f}$ is stable, we have

$$S_0 = S \cap \bigcap_{F \in \mathfrak{f}} {}^1V[F] \in \mathfrak{f}. \tag{1}$$

Let an arbitrary $c \in S_0$ be fixed. $\mathfrak{f}$ being free, there is a $T \in \mathfrak{f}$ such that $C \notin T \subset S$. Now $c \in {}^1V[T]$, so there is a $b \in T$ with $b\,{}^1V\,c \neq b$. As ${}^1Vb \cap \cap X \in \mathfrak{f}^1(b)$ and ${}^1V^{-1}c \cap X \in \mathfrak{f}^{-1}(c)$, the definition of ${}^1V$ gives that there are $x, y \in X$ such that $b\,{}^1V\,x\,V\,y\,{}^1V\,c$. From $V \subset {}^1V$ we obtain $x\,{}^1V\,y$, thus

$$b, c \in S, \quad y \in X, \quad b\,{}^1U\,y\,{}^1V\,c, \tag{2}$$

and so $U^*[S] \neq \emptyset$ (since ${}^1V \subset {}^1U$).

b) Let E denote the ${}^1\mathcal{U}$-envelope. $\mathfrak{f}^{\mathrm{E}} \subset \mathrm{fil}_Y\, \mathfrak{f}^*$ follows from $U^*[S] \subset {}^1U[S]$; thus $\mathfrak{f}^{\mathrm{E}} \subset \mathfrak{f}^{*\mathrm{E}}$. To prove the converse, take an $H \in \mathfrak{f}^{*\mathrm{E}}$. Then there are $S \in \mathfrak{f}$ and $U \in \mathcal{U}$ with ${}^1U[U^*[S]] \subset H$. We claim that if V and S_0 are as in a) (i.e. ${}^1V^2 \subset {}^1U$, and S_0 is defined by (1)) then ${}^1V[S_0] \subset {}^1U[U^*[S]]$, which implies that $H \in \mathfrak{f}^{\mathrm{E}}$. Let $d \in {}^1V[S_0]$, and pick a $c \in S_0$ such that $c\,{}^1V\,d$. Starting from an arbitrary $c \in S_0$, we obtained in a) points b and y satisfying (2). Now $y \in U^*[S]$ and $y\,{}^1V\,c\,{}^1V\,d$, therefore $d \in {}^1U[U^*[S]]$.

c) The statement on convergence is clear from b). Assume that a is a ${}^1\mathcal{U}^{\mathrm{tp}}$-cluster point of $\mathfrak{f}^*$; given $S \in \mathfrak{f}$ and $U \in \mathcal{U}$, we need a $c \in {}^1Ua \cap S$. Choose again $V \in \mathcal{U}$ with ${}^1V^2 \subset {}^1U$, and pick an $x \in {}^1Va \cap V^*[S]$. Now $x \in {}^1V^{-1}[S]$, so there is a $c \in S$ such that $x\,{}^1V\,c$; together with $a\,{}^1V\,x$, this implies $a\,{}^1U\,c$. □

10.3 LEMMA. *If $\mathfrak{P} \subset \mathfrak{P}_{\mathsf{S}}^{\mathrm{E}}$ overlaps $\mathfrak{P}_{\mathsf{S}}^{\mathrm{N}}$ then ${}^1\mathcal{U}(\mathfrak{P})$ is* SA*-complete.*

PROOF. Let $\mathfrak{f}^0$ be a stable Cauchy filter pair in ${}^1X$. If $\mathfrak{f}^0$ is not free, say there is an $a \in \bigcap \mathfrak{f}^{-1}$, then $\mathfrak{f}^1$ ${}^1\mathcal{U}^{\mathrm{tp}}$-converges to a by the Cauchy property, and clearly $\mathfrak{f}^{-1}$ ${}^1\mathcal{U}^{-\mathrm{tp}}$-clusters to a. So we can assume that $\mathfrak{f}^0$ is free. $\mathfrak{h}^0 = \mathfrak{f}^{0\mathrm{E}}|X$ is Cauchy, and it is also stable by Corollary 8.3. According to Lemma 10.2 b), $\mathfrak{h}^0$ is the $\mathcal{U}$-envelope of $\mathfrak{f}^{0*} = (\mathfrak{f}^{-1*}, \mathfrak{f}^{1*})$, so $\mathfrak{f}^{0*}$ is also stable and Cauchy. Now $\mathfrak{f}^{0*}$ has a cluster point a in ${}^1X$: if it has no cluster point in X then it has one in ${}^1X \setminus X$ because $\mathfrak{P}$ overlaps $\mathfrak{P}_{\mathsf{S}}^{\mathrm{N}}$. By Lemma 10.2 c), a is a cluster point of $\mathfrak{f}^0$. □

THEOREM. ${}^{\mathsf{SA}}\mathcal{U} = {}^1\mathcal{U}(\mathfrak{P}_{\mathsf{S}}^{\mathrm{NE}})$ *is a bitopological* SA*-completion; if $\mathcal{U}$ is a uniformity then ${}^{\mathsf{SA}}\mathcal{U}$ is its usual completion.* □

${}^{\mathsf{SA}}\mathcal{U}$ is neither a finest SA-completion, nor an SA-complete hull. We do not know whether there exists an SA-complete hull. On the other hand, the space in the example below has no finest SA-complete extension with stable trace filter pairs (it is not even finest among the SA-complete extensions with stable trace filter pairs; this can be of interest, because it would be possible to include into the definition of a completion that the trace filters or filter pairs belong to the class used when defining completeness).

EXAMPLE. Let $X = X_{-1} \cup X_1$, $X_i =]0, i[$, and $\mathcal{B} = \{U_\mathfrak{c} \colon \mathfrak{c} \in \mathfrak{C}\}$ a base for $\mathcal{U}$ on X, where $\mathfrak{C}$ consists of the countable partitions of X_1, and

$$x \, U_\mathfrak{c} \, y \text{ iff either } xy > 0, \; \exists C \in \mathfrak{c}, \; |x|, |y| \in C, \\ \text{or } x < 0 < y, \; -x, y \in Z_\mathfrak{c},$$

$$Z_\mathfrak{c} = \bigcup\{C \in \mathfrak{c} \colon C \text{ is uncountable}\}.$$

$U_\mathfrak{c}^2 \subset U_\mathfrak{c}$ and $U_{\mathfrak{c}(\cap)\mathfrak{d}} \subset U_\mathfrak{c} \cap U_\mathfrak{d}$, so $\mathcal{B}$ is indeed a base for a quasi-uniformity. $\mathcal{U}_1 = \mathcal{U} \mid X_1$ is a uniformity, which was considered in [I] II.50; it was shown there that the filter $\mathfrak{h}$ consisting of the co-countable subsets of X_1 is $\mathcal{U}_1$-stable, and no free filter is $\mathcal{U}_1$-Cauchy. Thus no free ultrafilter is $\mathcal{U}_1$-stable (e.g. by Lemma 5.8).

$\mathcal{U}$ induces the discrete bitopology. Let $\mathfrak{f}^0 \in \mathfrak{P}^{\mathsf{N}}_{\mathsf{C}}$. Then $\mathfrak{f}^0$ is free. $X_i \in \mathfrak{f}^i$ $(i = \pm 1)$, since if, say, $X_1 \in \sec \mathfrak{f}^{-1}$ then the Cauchy property implies that $X_1 \in \mathfrak{f}^1$, thus $\mathfrak{f}^0 \mid X_1$ is free and $\mathcal{U}_1$-Cauchy, implying that $\mathfrak{f}^1 \mid X_1$ is a free $\mathcal{U}_1$-Cauchy filter, a contradiction. Now from $X_i \in \mathfrak{f}^i$ and the Cauchy property we obtain that $\mathfrak{f}^0$ is finer than $\mathfrak{h}^0$ (which is Cauchy) where

$$\mathfrak{h}^1 = \operatorname{fil}_X \mathfrak{h}, \quad \mathfrak{h}^{-1} = \{-S \colon S \in \mathfrak{h}^1\}, \quad -S = \{-x \colon x \in S\}.$$

Thus $\mathfrak{P}^{\mathsf{N}}_{\mathsf{C}}$ consists of the filter pairs finer than $\mathfrak{h}^0$, and $\mathfrak{P}^{\mathsf{N}}_{\mathsf{S}}$ of the stable ones with this property.

Any $\mathfrak{f}^0 \in \mathfrak{P}^{\mathsf{N}}_{\mathsf{C}}$ is round, since if e.g. $S \in \mathfrak{f}^1$ and $S \subset X_1$ then $U_\mathfrak{c}[S] = S$ with $\mathfrak{c} = \{S, X_1 \setminus S\}$. Hence $\mathfrak{P}^{\mathsf{N}}_{\mathsf{S}} = \mathfrak{P}^{\mathsf{NE}}_{\mathsf{S}}$.

Assume that $\mathfrak{f}^0 \in \mathfrak{P}^{\mathsf{N}}_{\mathsf{S}}$. Then $\mathfrak{f}^1 \mid X_1$ is $\mathcal{U}_1$-stable, thus $\mathfrak{f}^1$ is not an ultrafilter. So there is an $A \subset X_1$ such that $A \in \sec \mathfrak{f}^1$ and $\mathfrak{e} = \operatorname{fil}_X (\mathfrak{f}^1 \mid A)$ is strictly finer than $\mathfrak{f}^1$. We claim that $\mathfrak{e}$ is $\mathcal{U}$-stable. Indeed, it is enough to consider partitions $\mathfrak{c}$ refining $\{A, X_1 \setminus A\}$ (since the entourages $U_\mathfrak{c}$ taken with such partitions form a base for $\mathcal{U}$). Now

$$\bigcap\{U_\mathfrak{c}[E] \colon E \in \mathfrak{e}\} = \bigcap\{U_\mathfrak{c}[E] \colon A \supset E \in \mathfrak{e}\} = \\ = \bigcap\{U_\mathfrak{c}[E \cup (X_1 \setminus A)] \cap A \colon A \supset E \in \mathfrak{e}\} \supset \bigcap\{U_\mathfrak{c}[S] \cap A \colon S \in \mathfrak{f}^1\} \in \mathfrak{f}^1 \mid A,$$

by the stability of $\mathfrak{f}^1$, hence $\mathfrak{e}$ is stable, too. This means that $(\mathfrak{f}^{-1}, \mathfrak{e}) \in \mathfrak{P}^{\mathsf{N}}_{\mathsf{S}}$, and it is strictly finer than $\mathfrak{f}^0$. Therefore $\mathfrak{P}^{\mathsf{N}}_{\mathsf{S}}$ has no maximal element.

Let $(Y, \mathcal{V})$ be an SA-complete extension with stable trace filter pairs. $\mathfrak{h}^0$ is a stable Cauchy filter base pair in Y, so it has a cluster point $p \in Y \setminus X$. Thus $\mathfrak{k}^0 = \mathfrak{f}^0(p)$ is stable Cauchy, and it overlaps $\mathfrak{h}^0$. Assume that some $x \in X$, say $x \in X_1$, is a cluster point of $\mathfrak{k}^0$. With $\mathfrak{c} = \{\{x\}, X_1 \setminus \{x\}\}$, take a $K \in \mathfrak{k}^\times$ such that $K \subset U_\mathfrak{c}$. $K_{-1} = K_1 = \{x\}$, contradicting that $\mathfrak{k}^0$ overlaps $\mathfrak{h}^0$. Therefore $\mathfrak{k}^0 \in \mathfrak{P}_\mathrm{S}^\mathrm{N}$. We claim that ${}^1\mathcal{U}(\mathfrak{P})$ is SA-complete, and the identity of X has no $(\mathcal{V}, {}^1\mathcal{U}(\mathfrak{P}))$-continuous extension, where $\mathfrak{P}$ consists of those elements of $\mathfrak{P}_\mathrm{S}^\mathrm{N} = \mathfrak{P}_\mathrm{S}^\mathrm{NE}$ that are not coarser than $\mathfrak{k}^0$; hence $\mathcal{V}$ cannot be a finest SA-complete extension. (It would be simpler to consider ${}^0\mathcal{U}(\mathfrak{P}_\mathrm{C}^\mathrm{Mn})$ instead of ${}^1\mathcal{U}(\mathfrak{P})$, but the elements of $\mathfrak{P}_\mathrm{C}^\mathrm{Mn}$ are not stable.)

$\mathfrak{P} \subset \mathfrak{P}_\mathrm{S}^\mathrm{E}$, and $\mathfrak{P}$ overlaps $\mathfrak{P}_\mathrm{S}^\mathrm{N}$, since if $\mathfrak{f}^0 \in \mathfrak{P}_\mathrm{S}^\mathrm{N}$ is not coarser than $\mathfrak{k}^0$ then $\mathfrak{f}^0 \in \mathfrak{P}$; otherwise, take a $\mathfrak{g}^0 \in \mathfrak{P}_\mathrm{S}^\mathrm{N}$ strictly finer than $\mathfrak{k}^0$, and then $\mathfrak{g}^0 \in \mathfrak{P}$ and $\mathfrak{g}^0$ is finer than $\mathfrak{f}^0$. Hence ${}^1\mathcal{U}(\mathfrak{P})$ is SA-complete by the lemma. The existence of a $(\mathcal{V}, {}^1\mathcal{U}(\mathfrak{P}))$-continuous extension f of the identity would imply that the trace filter pair of $f(p)$ is coarser than $\mathfrak{k}^0$, a contradiction. (We have seen that f cannot even be bitopologically continuous; the same could be said in some other examples, too.) □

§ 11 SF-completeness

11.1 Lemma 10.2 suggests the following definition: let us call a quasi-uniformity complete provided that each free stable Cauchy filter pair is convergent. This definition (yielding a notion between S- and SA-completeness) satisfies 2.1 a) to e), and ${}^1\mathcal{U}$ taken with the envelopes of the non-convergent free stable Cauchy filter pairs (or with a suitable subfamily of them) would be a complete extension having good properties. But there is an unwished-for consequence: as the example below shows, it would not depend only on the T_0 reflexion whether a quasi-uniformity is complete or not.

Example. On $X = A \cup B$ with $|A| = |B| = 2$, $A \cap B = \emptyset$, let $\mathcal{U} = \mathcal{U}(d)$ where $d(x, y) = 0$ if $x \in A$, $y \in B$. Each free (stable) Cauchy filter pair is convergent in $(X, \mathcal{U})$, but not in its product with an infinite indiscrete space: take the product of $(\mathrm{fil}_X\{A\}, \mathrm{fil}_X\{B\})$ and a free filter in the other space. □

Therefore we slightly modify the definition suggested above.

Definition. A filter $\mathfrak{f}$ in a quasi-uniform space $(X, \mathcal{U})$ is *fully free* if $\mathrm{cl}^s\{x\} \notin \sec \mathfrak{f}$ $(x \in X)$. A filter pair is *fully free* if both members of it are fully free. □

A filter (pair) is fully free iff its image in the T_0 reflexion is free.

11.2 Definition. A quasi-uniformity is SF-*complete* if each fully free stable Cauchy filter pair is convergent. □

A space is SF-complete iff its T_0 reflexion has the same property; a T_0 space is SF-complete iff each free stable Cauchy filter pair is convergent. S-completeness implies SF-completeness (evident), which implies SA-completeness (if $\mathfrak{f}^0$ is stable, Cauchy and, say, $\mathrm{cl}^s\{x\} \in \sec \mathfrak{f}^{-1}$ then x is a $\mathcal{U}^{-\mathrm{tp}}$-cluster point of $\mathfrak{f}^{-1}$; it is also a $\mathcal{U}^{\mathrm{tp}}$-cluster point of $\mathfrak{f}^{-1}$, so the Cauchy property gives that $\mathfrak{f}^1$ $\mathcal{U}^{\mathrm{tp}}$-converges to x).

REMARKS. a) It seems "more bitopological" to consider filter pairs $\mathfrak{f}^0$ for which $\mathrm{cl}^i\{x\} \notin \sec \mathfrak{f}^i$ $(i = \pm 1,\ x \in X)$. (Or perhaps $\mathrm{cl}^{-i}\{x\} \notin \sec \mathfrak{f}^i$.) Using either of these modified definitions, SF-completeness would not imply SA-completeness: $\mathcal{U}_{\mathrm{se}} \mid \mathsf{R}_0$ would be SF-complete.

b) Neighbourhood filter pairs are, of course, not (fully) free; nevertheless, SF-completeness can be regarded as a bitopological notion in the sense of 1.1, cf. "To K" in 2.2.

LEMMA. *SF-completeness satisfies Conditions 2.1 a) to e).*

PROOF. a) is clear. b), d) and e) hold for S- and SA-completeness, hence for any notion between them. c) follows from the observation that a fully free filter given in a subspace generates a fully free filter in the whole space: if $\mathfrak{f}$ is a filter in $S \subset X$, $x \in X$ and $\mathrm{cl}^s\{x\} \in \sec \mathfrak{f}$ then there is an $y \in \mathrm{cl}^s\{x\} \cap S$, and $\mathrm{cl}^s_S\{y\} = \mathrm{cl}^s\{y\} \cap S = \mathrm{cl}^s\{x\} \cap S \in \sec_S \mathfrak{f}$. □

11.3 NOTATION. In $(X,\mathcal{U})$, $\mathfrak{P}_{\mathsf{F}} = \mathfrak{P}_{\mathsf{F}}(\mathcal{U})$ is the family of the non-convergent fully free stable Cauchy filter pairs. □

LEMMA. *$\mathfrak{P}_{\mathsf{F}}$ is overlayed by $\mathfrak{P}^{\mathrm{m}}_{\mathsf{F}}$.*

PROOF. Copy the proof of Lemma 5.5, using that the intersection of fully free filters has the same property. □

In contrast to $\mathfrak{P}^{\mathrm{m}}_{\mathsf{S}}$, the elements of $\mathfrak{P}^{\mathrm{m}}_{\mathsf{F}}$ are not always round: on $X = \mathsf{R}$, let $d(x,y) = d_{\mathrm{so}}(x,y)$ if $xy \neq 0$, and $d(x,0) = 0$; now $\mathfrak{f}^0 = \mathrm{Fil}\ (\mathfrak{e}_0 \mid \mathsf{R}_0) \in \mathfrak{P}^{\mathrm{m}}_{\mathsf{F}}$, but $\mathfrak{f}^{0\mathrm{E}}$ is not free.

11.4 LEMMA. *If $\mathfrak{f}^0, \mathfrak{g}^0 \in \mathfrak{P}^{\mathrm{m}}_{\mathsf{F}}$ and $\mathfrak{g}^{0\mathrm{E}}$ is finer than $\mathfrak{f}^{0\mathrm{E}}$ then $\mathfrak{f}^0 = \mathfrak{g}^0$.*

PROOF. $\mathfrak{h}^0 = \mathfrak{f}^0 \cap \mathfrak{g}^0$ is finer than the Cauchy filter pair $\mathfrak{f}^{0\mathrm{E}}$, thus it is Cauchy, too. $\mathfrak{h}^0$ is also stable and fully free, so $\mathfrak{h}^0 \in \mathfrak{P}_{\mathsf{F}}$, implying $\mathfrak{h}^0 = \mathfrak{f}^0 = \mathfrak{g}^0$. □

11.5 NOTATION. In $(X,\mathcal{U})$, $\mathfrak{F}_{\mathsf{F}}(\mathcal{U})$ is the family of the fully free stable filters. □

LEMMA. *If ${}^1\mathcal{U}$ is taken with trace filter pairs belonging to $\mathfrak{P}^{\mathrm{E}}_{\mathsf{F}}$, and $\mathfrak{f} \in \mathfrak{F}^{\mathrm{E}}_{\mathsf{F}}({}^1\mathcal{U})$, then $\mathfrak{f} \mid X \in \mathfrak{F}^{\mathrm{E}}_{\mathsf{F}}(\mathcal{U})$.*

PROOF. Let $\mathfrak{f} \in \mathfrak{F}^{\mathrm{E}}_{\mathsf{F}}({}^1\mathcal{U})$ be the ${}^1\mathcal{U}$-envelope of the fully free stable filter $\mathfrak{h}$, and $\mathfrak{f}^1(p)$ the $\mathcal{U}$-envelope of the fully free stable filter $\mathfrak{h}^1(p)$. A filter in X

is defined by

(1) $\quad \mathfrak{k}=\{(S\cap X)\cup\bigcup\{H(p): p\in S\setminus X\}: S\in\mathfrak{h}, H(p)\in\mathfrak{h}^1(p)\ (p\in S\setminus X)\}.$

$\mathfrak{k}$ is clearly fully free. We are going to show that $\mathfrak{f}\,|\,X=\mathfrak{k}^{\mathrm{E}}$; this will also imply that $\mathfrak{k}$ is stable, since the stability of $\mathfrak{k}^{\mathrm{E}}$ follows from Corollary 8.3.

If $F\in\mathfrak{f}\,|\,X$ then there are $U\in\mathcal{U}$ and $S\in\mathfrak{h}$ such that $F={}^1U[S]\cap X$. The union of $S\cap X$ and of the sets $H(p)={}^1Up\cap X\in\mathfrak{f}^1(p)\subset\mathfrak{h}^1(p)$ $(p\in S\setminus X)$ is a subset of F belonging to $\mathfrak{k}$ by definition. Thus $\mathfrak{f}\,|\,X\subset\mathfrak{k}$, implying $\mathfrak{f}\,|\,X\subset\mathfrak{k}^{\mathrm{E}}$, since $\mathfrak{f}\,|\,X$ is round.

Conversely, let now $F\in\mathfrak{k}^{\mathrm{E}}$, and pick $U\in\mathcal{U}$, $S\in\mathfrak{h}$ and $H(p)\in\mathfrak{h}^1(p)$ $(p\in S\setminus X)$ such that $F=U[T]$ where $T=(S\cap X)\cup\bigcup\{H(p): p\in S\setminus X\}$. We claim that if $V\in\mathcal{U}$ with $V^3\subset U$ then ${}^1V[S]\cap X\subset F$; this means that $F\in\mathfrak{f}\,|\,X$.

Indeed, take a point $z\in{}^1V[S]\cap X$, and pick an $a\in S$ such that $a\,{}^1V\,z$. Let $A=Va$ if $a\in X$, $A=V[H(a)]$ if $a\notin X$, $B=V^{-1}z$. As $A\in\mathfrak{f}^1(a)$ and $B\in\mathfrak{f}^{-1}(z)$, there are $x\in A$ and $y\in B$ with $x\,V\,y$ (even if $a=z$, since then $x=y=z$ will do). If $a\in X$ then $a\,V^3\,z$, so $a\,U\,z$ and $z\in U[S\cap X]\subset U[T]=$ $=F$; if $a\notin X$ then there is a $w\in H(a)$ such that $w\,V\,x$; now $w\,U\,z$ and $z\in U[H(a)]\subset U[T]=F$. $\square$

Remark. Observe for later use that if each of the filters $\mathfrak{h}$ and $\mathfrak{h}^1(p)$ $(p\notin X)$ is an ultrafilter then so is $\mathfrak{k}$: Let $X=A\cup B$. If $X\in\mathfrak{h}$ then $A\in\mathfrak{h}$ or $B\in\mathfrak{h}$, so it is clear from (1) that $A\in\mathfrak{k}$ or $B\in\mathfrak{k}$. Thus we can assume that ${}^1X\setminus X\in\mathfrak{h}$. Define

$$A'=\{p\in{}^1X\setminus X: A\in\mathfrak{h}^1(p)\},$$

and B' analogously. Each $\mathfrak{h}^1(p)$ being an ultrafilter, we have $A'\cup B'={}^1X\setminus X$, so, say, $A'\in\mathfrak{h}$. Now with $S=A'$ and $H(p)=A$ $(p\in S)$, (1) gives $A\in\mathfrak{k}$.

11.6 Notation. ${}^{\mathsf{SF}}\mathcal{U}={}^1\mathcal{U}(\mathfrak{P}_{\mathsf{F}}^{\mathrm{mE}})$. $\square$

Theorem. *${}^{\mathsf{SF}}\mathcal{U}$ is a basic bitopological SF-completion of $\mathcal{U}$; it is finest, and a complete hull; for uniformities, it coincides with the usual completion.*

Proof. The elements of $\mathfrak{P}_{\mathsf{F}}^{\mathrm{mE}}$ are round stable and Cauchy, so ${}^{\mathsf{SF}}\mathcal{U}$ is an extension of $\mathcal{U}$. To show that ${}^{\mathsf{SF}}\mathcal{U}$ is SF-complete, take a fully free stable Cauchy filter pair $\mathfrak{h}^0$ in it, and let $\mathfrak{f}^0$ be the ${}^{\mathsf{SF}}\mathcal{U}$-envelope of $\mathfrak{h}^0$. Lemma 11.5 applied to $\mathcal{U}$ as well as to $\mathcal{U}^{-1}$ gives that there is a fully free stable filter pair $\mathfrak{k}^0$ in $(X,\mathcal{U})$ such that $\mathfrak{f}^0\,|\,X=\mathfrak{k}^{0\mathrm{E}}$. Moreover, $\mathfrak{k}^0$ is Cauchy since $\mathfrak{f}^0$ is ${}^1\mathcal{U}$-Cauchy. If $\mathfrak{k}^0$ is not convergent then $\mathfrak{k}^0\in\mathfrak{P}_{\mathsf{F}}$, thus Lemma 11.3 implies that $\mathfrak{k}^0$ is finer than some element of $\mathfrak{P}_{\mathsf{F}}^{\mathrm{mE}}$. So the filter base pair $\mathfrak{k}^0$ in $({}^1X,{}^1\mathcal{U})$ (hence $\mathfrak{f}^0\,|\,X$, too) converges in any case to some $a\in{}^1X$. By Lemma 10.2 b), $(\mathfrak{h}^{-1*},\mathfrak{h}^{1*})^{\mathrm{E}}=\mathfrak{f}^0\,|\,X$, so Lemma 10.2 c) implies that $\mathfrak{h}^0$ converges to a.

${}^{\mathsf{SF}}\mathcal{U}$ is clearly basic and bitopological. To prove that it is finest, take another SF-complete extension $(Y,\mathcal{V})$. For $p\in{}^{\mathsf{SF}}X\setminus X$, $\mathfrak{f}^0(p)$ is the envelope of some $\mathfrak{h}^0(p) \in \mathfrak{P}^{\mathrm{m}}_{\mathsf{F}}$, so $\mathfrak{f}^0(p)$ $\mathcal{V}$-converges to some $f(p) \in Y$. Letting $f(x) = x$ $(x \in X)$, we obtain an $({}^{\mathsf{SF}}\mathcal{U}, \mathcal{V})$-continuous map f, since if $V \in \mathcal{V}$ and $U = V \mid X$ then $a\,{}^1U\,b$ implies $f(a)\,V^3\,f(b)$. Indeed, $A = V f(a) \cap X \in \mathfrak{f}^1(a)$ and $B = V^{-1} f(b) \cap X \in \mathfrak{f}^{-1}(b)$ (by the $\mathcal{V}$-convergence of the ${}^{\mathsf{SF}}\mathcal{U}$-trace filter pairs), so if $a\,{}^1U\,b$ and $a \neq b$ then there are $x \in A$ and $y \in B$ with $x\,U\,y$, i.e. $x\,V\,y$ and $f(a)\,V^3\,f(b)$.

Assume that a subspace Z of $({}^{\mathsf{SF}}X, {}^{\mathsf{SF}}\mathcal{U})$ is an SF-complete extension of $\mathcal{U}$. Given $\mathfrak{g}^0 \in \mathfrak{P}^{\mathrm{m}}_{\mathsf{F}}(\mathcal{U})$, there has to be a $p \in Z \setminus X$ to which $\mathfrak{g}^0$ (and so $\mathfrak{g}^{0\mathrm{E}}$) ${}^{\mathsf{SF}}\mathcal{U}^{\mathrm{tp}} \mid Z$-converges, thus $\mathfrak{g}^{0\mathrm{E}}$ is finer than $\mathfrak{f}^0(p)$, which is of the form $\mathfrak{f}^{0\mathrm{E}}$ where $\mathfrak{f}^0 \in \mathfrak{P}^{\mathrm{m}}_{\mathsf{F}}$. Now Lemma 11.4 implies that $\mathfrak{f}^0 = \mathfrak{g}^0$, so the point in ${}^{\mathsf{SF}}X$ belonging to the trace filter pair $\mathfrak{g}^{0\mathrm{E}}$ is in Z, therefore $Z = {}^{\mathsf{SF}}X$, i.e. ${}^{\mathsf{SF}}\mathcal{U}$ is an SF-complete hull.

Assume finally that $\mathcal{U}$ is a uniformity. Then ${}^{\mathsf{SF}}\mathcal{U}$ is also a uniformity, and it is a complete hull by Lemma 11.2 and the preceding paragraph. Hence ${}^{\mathsf{SF}}\mathcal{U}$ is the usual completion. □

Neither the extension theorem for maps nor i″) is valid:

Examples. a) With $(X^{\circ\circ}, \mathcal{U}^{\circ\circ})$ from Example 6.6 and $(X, \mathcal{U})$ from Example 11.1, $\mathcal{U}$ is SF-complete, and one can easily define a uniformly continuous map f from $X^{\circ\circ}$ onto X, shrinking the intervals in $X^{\circ\circ}$ to single points. But f has no uniformly continuous extension to ${}^{\mathsf{SF}}(X^{\circ\circ})$: there is nowhere to map the point with the trace filter pair 6.6 (1).

b) In the space from Example 6.4, $\mathfrak{P}^{\mathrm{mE}}_{\mathsf{F}}$ consists of just the filter pairs considered there, and ${}^5\mathcal{U}$ is an SF-complete extension strictly coarser than ${}^{\mathsf{SF}}\mathcal{U}$. □

§ 12 U-completeness

12.1 A space is clearly A-complete iff each Cauchy ultrafilter pair is convergent. The analogous statement for SA-completeness is false: $\mathcal{U}$ from Example 10.3 is not SA-complete, but each stable Cauchy ultrafilter pair is convergent (because the stable ultrafilters are fixed). Thus the following definition gives a notion strictly weaker than SA-completeness:

Definition. A quasi-uniformity is U-*complete* if each stable Cauchy ultrafilter pair is convergent. □

A space is U-complete iff those stable Cauchy filter pairs are convergent that are envelopes of ultrafilter pairs; this way we can formulate the definition with a class containing the neighbourhood filter pairs. U-completeness implies L-completeness: If $\mathfrak{f}^0$ is linked and Cauchy then there is an ultrafilter $\mathfrak{h}$ finer than $\mathfrak{f}^{-1}(\cap)\mathfrak{f}^1$. Now $(\mathfrak{h}, \mathfrak{h})$ is a Cauchy ultrafilter pair, which is also stable (because linked Cauchy). Then $\mathfrak{f}^0$ clusters to the limit points of $(\mathfrak{h}, \mathfrak{h})$.

(And this is sufficient for the L-completeness, see in 4.1.) U-completeness satisfies all the conditions in 1.1: b) and e) because it is between SA- and L-completeness; a) is evident; c) and d) are easy to check.

12.2 NOTATION. $\mathfrak{P}_{\mathrm{U}}$ is the class of the non-convergent stable Cauchy ultrafilter pairs. □

The elements of $\mathfrak{P}_{\mathrm{U}}$ are clearly fully free. Different elements of $\mathfrak{P}_{\mathrm{U}}$ can have the same envelope; but $\mathfrak{P}_{\mathrm{U}}^{\mathrm{E}}$ is to be understood as a set, i.e. its elements are not taken with multiplicity.

LEMMA. *$\mathfrak{P}_{\mathrm{U}}^{\mathrm{E}}$ is overlayed by $\mathfrak{P}_{\mathrm{U}}^{\mathrm{Em}}$.*

PROOF. Given $\mathfrak{f}^0 \in \mathfrak{P}_{\mathrm{U}}^{\mathrm{E}}$, let $\mathfrak{f} \in \mathfrak{F}$ iff $\mathfrak{f} \subset \mathfrak{f}^1$ and $(\mathfrak{f}^{-1}, \mathfrak{f}) \in \mathfrak{P}_{\mathrm{U}}^{\mathrm{E}}$. We are going to show that Zorn's Lemma can be applied to $\mathfrak{F}$ taken with the reverse inclusion.

Take $\mathfrak{f}_\alpha \in \mathfrak{F}$ for $\alpha \in I$ where I is linearly ordered and $\mathfrak{f}_\alpha \supset \mathfrak{f}_\beta$ whenever $\alpha < \beta$. For each $\alpha \in I$, pick an ultrafilter $\mathfrak{h}_\alpha$ such that $\mathfrak{f}_\alpha = \mathfrak{h}_\alpha^{\mathrm{E}}$. Let $\mathfrak{i}$ be an ultrafilter in I consisting of cofinal sets. Define

$$\mathfrak{h} = \{S : \exists I_0 \in \mathfrak{i},\ S \in \mathfrak{h}_\alpha\ (\alpha \in I_0)\}. \tag{1}$$

$\mathfrak{h}$ is clearly a filter. Assuming $A \cup B = X$, let $A' = \{\alpha \in I : A \in \mathfrak{h}_\alpha\}$ and $B' = \{\alpha \in I : B \in \mathfrak{h}_\alpha\}$. Now $A' \cup B' = I$, thus $A' \in \mathfrak{i}$ or $B' \in \mathfrak{i}$, implying that $A \in \mathfrak{h}$ or $B \in \mathfrak{h}$. Hence $\mathfrak{h}$ is an ultrafilter. Given $\alpha_0 \in I$, $S \in \mathfrak{h}$ and $U \in \mathcal{U}$, choose I_0 according to (1), and $\alpha > \alpha_0$ such that $\alpha \in I_0$. Now $S \in \mathfrak{h}_\alpha$, so $U[S] \in \mathfrak{h}_\alpha^{\mathrm{E}} = \mathfrak{f}_\alpha \subset \mathfrak{f}_{\alpha_0}$ i.e. $\mathfrak{h}^{\mathrm{E}} \subset \mathfrak{f}_{\alpha_0}$ for each $\alpha_0 \in I$. Conversely, any element of $\mathfrak{g} = \bigcap_{\alpha \in I} \mathfrak{f}_\alpha$ belongs to each $\mathfrak{h}_\alpha$, thus to $\mathfrak{h}$ by (1). $\mathfrak{h}^{\mathrm{E}} = \mathfrak{g}^{\mathrm{E}}$ is a filter coarser than each $\mathfrak{f}_\alpha$; it is stable by Lemma 5.4; it is clearly non-convergent; $(\mathfrak{f}^{-1}, \mathfrak{h}^{\mathrm{E}})$ is Cauchy, because $(\mathfrak{f}^{-1}, \mathfrak{g})$ is Cauchy by Lemma 5.2. Thus $\mathfrak{h}^{\mathrm{E}} \in \mathfrak{F}$.

Take a minimal filter $\mathfrak{k}^1 \subset \mathfrak{f}^1$ for which $(\mathfrak{f}^{-1}, \mathfrak{k}^1) \in \mathfrak{P}_{\mathrm{U}}^{\mathrm{E}}$, and then similarly a minimal filter $\mathfrak{k}^{-1} \subset \mathfrak{f}^{-1}$ for which $\mathfrak{k}^0 \in \mathfrak{P}_{\mathrm{U}}^{\mathrm{E}}$. Now $\mathfrak{k}^0$ is coarser than $\mathfrak{f}^0$, and $\mathfrak{k}^0 \in \mathfrak{P}_{\mathrm{U}}^{\mathrm{Em}}$. □

12.3 NOTATION. ${}^{\mathrm{U}}\mathcal{U} = {}^{1}\mathcal{U}(\mathfrak{P}_{\mathrm{U}}^{\mathrm{Em}})$. □

THEOREM. *${}^{\mathrm{U}}\mathcal{U}$ is a basic bitopological* U*-completion; it is a complete hull; for uniformities, it coincides with the usual completion.*

PROOF. To prove that ${}^{\mathrm{U}}\mathcal{U}$ is U-complete, take a stable Cauchy ultrafilter pair $\mathfrak{h}^0$ in ${}^{\mathrm{U}}X$, and let $\mathfrak{f}^0$ be its ${}^{\mathrm{U}}\mathcal{U}$-envelope; we may assume that $\mathfrak{h}^0$ is fully free. Lemma and Remark 11.5 yield a stable Cauchy ultrafilter pair $\mathfrak{k}^0$ such that $\mathfrak{f}^0 \mid X = \mathfrak{k}^{0\mathrm{E}}$. Thus $\mathfrak{f}^0 \mid X$ is finer than a trace filter pair, $\mathfrak{f}^0$ converges to the corresponding point by Lemma 10.2 (applied in the same way as in the proof of Theorem 11.6).

To complete the proof, copy the corresponding parts from the proof of Theorem 11.6. □

12.4 Similarly to SF-completeness, U-completeness fulfils all the requirements a) to i) from § 1. In fact, U-completeness is even better:

PROPOSITION. *The extension theorem for maps holds for the* U*-completion.*

PROOF. Let f be a uniformly continuous map from $(X,\mathcal{U})$ into a U-complete space $(Y,\mathcal{V})$. For $p \in {}^{\mathsf{U}}X \setminus X$, let $\mathfrak{h}^0$ be an ultrafilter pair such that $\mathfrak{f}^0(p) = \mathfrak{h}^{0\mathrm{E}}$. Now $(f(\mathfrak{h}^{-1}), f(\mathfrak{h}^1))$ is a stable Cauchy ultrafilter pair in Y, so we can pick a point $f(p)$ to which it converges. The extended function f obtained this way is $({}^{\mathsf{U}}\mathcal{U}, \mathcal{V})$-continuous: Given $V \in \mathcal{V}$, choose $U \in \mathcal{U}$ such that $f(x)\, V\, f(y)$ whenever $x\, U\, y$. Now $a\, {}^1U\, b$ implies $f(a)\, V^3\, f(b)$. □

REMARK. This argument cannot be used for the SF-completion, since the image of a fully free filter may not be fully free.

§ 13 R-completeness

13.1 We are going to introduce a bitopological notion on the analogy of W-completeness (see in 2.1). The properties of this notion are not as good as those of SF- or U-completeness. On the other hand, it will be very easy to obtain a good completion, just copying the construction of ${}^{\mathsf{W}}\mathcal{U}$ (see in 2.2).

DEFINITION. A quasi-uniformity is

a) R-*complete* if each round Cauchy filter pair has a cluster point;

b) SR-*complete* if each stable round Cauchy filter pair has a cluster point. □

Both notions satisfy 1.1 a), b) and d), but neither c) nor e). To c): we shall return to it in Example 13.2. To e): in $(\mathbb{R}_0, \mathcal{U}_{\mathrm{se}} \mid \mathbb{R}_0)$, any point is a cluster point of any round filter pair, thus $\mathcal{U}_{\mathrm{se}} \mid \mathbb{R}_0$ is R-complete.

13.2 Before defining the R-completion, let us first observe that for any round filter pair $\mathfrak{f}^0$ there is a maximal round filter pair finer than it, which is clearly Cauchy if $\mathfrak{f}^0$ was so. (Zorn's Lemma, as in [Cs5] 1.4.) Hence a space is R-complete iff each maximal round Cauchy filter pair is convergent. (Because if a maximal round filter (pair) clusters to x then it converges to x, cf. [Cs5] 1.6.)

NOTATION. ${}^{\mathsf{R}}\mathcal{U} = {}^{0}\mathcal{U}(\mathfrak{P}^{\mathrm{Mn}}_{\mathsf{R}})$. □

THEOREM. ${}^{\mathsf{R}}\mathcal{U}$ *is a bitopological* R*-completion; it is finest and a complete hull; it is a unique finest basic* R*-completion; for uniformities, it coincides with the usual completion.*

PROOF. Let $\mathfrak{f}^0$ be a round Cauchy filter pair in ${}^{\mathsf{R}}X$. Then $\mathfrak{f}^0 \mid X$ is round and Cauchy in X, so there is an $\mathfrak{h}^0 \in \mathfrak{P}^{\mathrm{M}}_{\mathsf{R}}$ finer than it. Now if $\mathfrak{h}^0$ is not convergent then it is a trace filter pair; thus the filter base pair $\mathfrak{h}^0$ in

${}^{\mathsf{R}}X$ is convergent, therefore $\mathfrak{f}^0$ has a cluster point. So ${}^{\mathsf{R}}\mathcal{U}$ is R-complete; it is evidently basic and bitopological.

Assume that with some $Z \subset {}^{\mathsf{R}}X$, ${}^{\mathsf{R}}\mathcal{U} \mid Z$ is an R-complete extension, and take a $p \in {}^{\mathsf{R}}X \setminus X$. Let $\mathfrak{f}^0$ be the ${}^{\mathsf{R}}\mathcal{U} \mid Z$-envelope of $\mathfrak{f}^0(p)$, and q a cluster point of $\mathfrak{f}^0$; clearly, $q \in Z \setminus X$. Assume that $p \neq q$. Then $\mathfrak{f}^0(p) \neq \mathfrak{f}^0(q)$, say, $\mathfrak{f}^1(p) \neq \neq \mathfrak{f}^1(q)$. As $\mathfrak{f}^1(p)$ and $\mathfrak{f}^1(q)$ are maximal round filters, there are disjoint sets $A \in \mathfrak{f}^1(p)$, $B \in \mathfrak{f}^1(q)$ ([Cs5] 1.5). Take $V \in {}^{\mathsf{R}}\mathcal{U}$ and $S \in \mathfrak{f}^1(p)$ such that $V^2 q \cap \cap X \subset B$ and $V^2[S] \cap X \subset A$. Since q is a cluster point of the ${}^{\mathsf{R}}\mathcal{U}$-envelope of $\mathfrak{f}^1(p)$, there is an $a \in V[S] \cap Vq$. Now $Va \cap X \neq \emptyset$, hence $A \cap B \neq \emptyset$, a contradiction. Thus $p = q$, $p \in Z$ and $Z = {}^{\mathsf{R}}X$, i.e. ${}^{\mathsf{R}}\mathcal{U}$ is an R-complete hull.

To prove that ${}^{\mathsf{R}}\mathcal{U}$ is finest, take another R-complete extension $(Y, \mathfrak{V})$. Define $f(x) = x$ $(x \in X)$, and, for $p \in {}^{\mathsf{R}}X \setminus X$, pick a cluster point $f(p) \in Y$ of the $\mathcal{V}$-envelope of $\mathfrak{f}^0(p)$. Just as in the preceding paragraph, double density implies that $\mathfrak{f}^0(p)$ itself clusters to $f(p)$. Since $\mathfrak{f}^0(p)$ is a maximal round filter pair, the $\mathcal{V}$-trace filter pair of $f(p)$ is coarser than $\mathfrak{f}^0(p)$. Given $W \in \mathcal{V}$, let $g^i(a) = W^i f(a) \cap X$ (which is in $\mathfrak{f}^i(a)$); now $a\, V(g^{-1}, g^1, W \mid X)\, b$ implies $f(a)\, W^3\, f(b)$. Thus f is uniformly continuous, and ${}^{\mathsf{R}}\mathcal{U}$ is finest.

Assume now that $(Y, \mathcal{V})$ is another finest basic R-complete extension, and let f be a $(\mathcal{V}, {}^{\mathsf{R}}\mathcal{U})$ continuous extension of the identity of X. By the continuity, the $\mathcal{V}$-trace filter pair $\mathfrak{h}^0(p)$ of $p \in Y \setminus X$ has to be finer than $\mathfrak{f}^0(f(p))$; here $f(p) \notin X$, since $\mathcal{V}$ is basic, and so the trace filter pairs of the new points are not convergent. $\mathfrak{f}^0(f(p))$ is maximal round, thus $\mathfrak{h}^0(p) = = \mathfrak{f}^0(f(p))$. Different new points have different trace filter pairs in basic extensions, so f is injective. Identifying Y with $f[Y]$, we obtain that $\mathcal{V}$ is an extension finer than ${}^{\mathsf{R}}\mathcal{U} \mid Y$ such that they induce the same trace filter pairs; but ${}^{\mathsf{R}}\mathcal{U} \mid Y$ was already the finest one (of the form ${}^{0}\mathcal{U}$), thus $\mathcal{V} = {}^{\mathsf{R}}\mathcal{U} \mid Y$. This implies $\mathcal{V} = {}^{\mathsf{R}}\mathcal{U}$, since ${}^{\mathsf{R}}\mathcal{U}$ is a complete hull. Hence ${}^{\mathsf{R}}\mathcal{U}$ is indeed a unique finest basic R-completion.

If $\mathcal{U}$ is a uniformity then ${}^{\mathsf{R}}\mathcal{U}$ is its usual completion, since (i) ${}^{\mathsf{R}}\mathcal{U}$ is a uniformity, (ii) R-completeness is the usual completeness for uniformities, (iii) it is a complete hull. □

The next example shows that 1.2 i′) is not valid for the R-completion.

Example. Just as in 2.2 "To W", let $X =]0,1] \times \{0\}$, $X_* =]0,1] \times [0,1]$, $\mathcal{U}_* = \mathcal{U}(d_{\mathrm{eu}} \times d_0) \mid X_*$, $\mathcal{U} = \mathcal{U}_* \mid X$, where $d_0(x,y) = |y - x|$ if $y \neq 0$. Now $(X, \mathcal{U})$ is a subspace of the R-complete space $(Y, \mathcal{V}) = ({}^{\mathsf{R}}X_*, {}^{\mathsf{R}}\mathcal{U}_*)$, but the identity of X has no $({}^{\mathsf{R}}\mathcal{U}, \mathcal{V})$-continuous extension. Indeed, $\mathcal{U}$ is the Euclidean uniformity of the half-closed interval, so there is a single point $p \in {}^{\mathsf{R}}X \setminus X$, with $\mathfrak{f}^0(p) = (\mathfrak{e}_1^2, \mathfrak{e}_1^2) \mid X$. If there existed a uniformly continuous extension then we would obtain a $q \in Y \setminus X$ such that the $\mathcal{V}$-trace filter pair $\mathfrak{h}^0$ of q (on X_*) is coarser than $\mathfrak{f}^0(p)$, hence than the $\mathcal{U}_*$-envelope $\mathfrak{g}^0$ of $\mathfrak{f}^0(p)$. But this is impossible, since $\mathfrak{h}^0$ is a maximal round filter pair, while $\mathfrak{g}^0$ is not. □

In the above example, Y can be identified with $[0,1]^2$, such that $\mathcal{V}=\mathcal{U}(d)$ where $d(x,y)=d_{\mathrm{eu}}(x,y)$ if $y\notin X$. Thus $(X,\mathcal{U})$ is a quasi-closed (because $\mathcal{V}^{\mathrm{tp}}$-closed) subspace of $(Y,\mathcal{V})$. Hence 1.1 c) does not hold for R-completeness; $\mathcal{V}$ being totally bounded, all the filter pairs are stable, thus 1.1 c) cannot hold for SR-completeness either.

13.3 We cannot obtain an SR-completion just by considering only stable filter pairs in the construction of ${}^{\mathsf{R}}\mathcal{U}$, since for a round stable Cauchy filter pair, there is in general no maximal round stable one finer than it: in Example 10.3, $\mathfrak{P}_{\mathsf{S}}^{\mathrm{N}}\neq\emptyset$ has no maximal element, and each element of it is round, hence $\mathfrak{P}_{\mathsf{S}}^{\mathrm{EN}}\neq\emptyset$ has no maximal element either. ${}^{\mathsf{SR}}\mathcal{U}={}^{1}\mathcal{U}(\mathfrak{P}_{\mathsf{S}}^{\mathrm{EN}})$ is an SR-completion (using Corollary 8.3, the proof is straightforward). ${}^{\mathsf{SR}}\mathcal{U}$ is neither finest, nor a complete hull (again Example 10.3).

§ 14 Comparing the bitopological notions of completeness

14.1 In addition to the notions introduced so far, the following definitions seem also reasonable: a quasi-uniformity is F-*complete* if each fully free Cauchy filter pair is convergent; it is O-*complete* (SO-*complete*, LO-*complete*, LR-*complete*) if each open (open stable, linked open, linked round) Cauchy filter pair has a cluster point. We ignore some more complicated possibilities: the envelope of each stable ultrafilter pair has a cluster point; each stable maximal open Cauchy filter pair is convergent; etc. Some of the implications in the diagram below have already been proved, the others are clear from the definitions.

$$\begin{tikzcd}
\mathsf{C} \arrow[r, Rightarrow] \arrow[d, Rightarrow] & \mathsf{F} \arrow[r, Rightarrow] \arrow[d, Rightarrow] & \mathsf{A} \arrow[r, Rightarrow] \arrow[d, Rightarrow] & \mathsf{O} \arrow[rr, Rightarrow] \arrow[d, Rightarrow] & & \mathsf{R} \arrow[d, Rightarrow] & \\
\mathsf{S} \arrow[r, Rightarrow] & \mathsf{SF} \arrow[r, Rightarrow] & \mathsf{SA} \arrow[r, Rightarrow] \arrow[d, Rightarrow] & \mathsf{SO} \arrow[rr, Rightarrow] \arrow[ddr, Rightarrow] & & \mathsf{SR} \arrow[ddr, Rightarrow] & \\
& & \mathsf{U} \arrow[d, Rightarrow] & & & & \\
& & \mathsf{L} \arrow[rr, Rightarrow] & & \mathsf{LO} \arrow[rr, Rightarrow] & & \mathsf{LR}
\end{tikzcd}$$

The properties in the same column coincide in totally bounded spaces: $\mathsf{L}\Rightarrow\mathsf{A}$, because if $\mathfrak{f}^0$ is a Cauchy filter pair then, with an arbitrary ultrafilter $\mathfrak{h}$ finer than $\mathfrak{f}^{-1}$, $(\mathfrak{h},\mathfrak{h})$ is Cauchy, and so is $(\mathfrak{h},\mathfrak{h}\cap\mathfrak{f}^1)$; now any limit point of this linked Cauchy filter pair is a cluster point of $\mathfrak{f}^0$. The other implications follow from Corollary 5.8. If T_1 is assumed in addition to total boundedness then $\mathsf{L}\Rightarrow\mathsf{C}$, since, according to [LF] Corollary 21, an L-complete totally bounded T_1 quasi-uniformity is a uniformity.

14.2 The examples below show that no more implications are valid between the notions in the diagram. We shall try to give quasi-(pseudo)metrizable examples, such that the failure of a completeness property is guaranteed by a filter pair generated by a sequence pair (because the sequential version of some of the properties could perhaps also be of interest; cf. [Kü] for sequential t o p o l o g i c a l notions of quasi-pseudometric completeness). Whenever possible, the examples will be totally bounded and/or T_1.

EXAMPLES. a) F, S *not* C. Let $X=\mathbb{R}_1\times\{0\}\cup\{(0,-1),(0,1)\}, \mathcal{U}=\mathcal{U}(d)$,

$$d(x,y)=y' \quad \text{if} \quad x'=0<y'.$$

A Cauchy filter pair $\mathfrak{f}^0$ is non-convergent iff $\mathfrak{f}^{-1}=\operatorname{fil}\{\{(0,-1),(0,1)\}\}$ and $\mathfrak{f}^1$ is finer than $\mathfrak{e}_1^2\mid X$. These filter pairs are neither stable nor free.

b) A, S, *not* F. On $X=\mathbb{R}_{-1}\times\{0\}\cup\overline{\mathbb{R}}_1\times\{-1,1\}$, let $\mathcal{U}=\mathcal{U}(d)$ with

$$d(x,y)=y'-x' \quad \text{if either } x'<0\leqq y', \qquad \text{or } x'=0<y',\ x''=y''.$$

The non-convergent Cauchy filter pairs cluster to $(0,-1)$ and $(0,1)$; a filter pair clustering to both points cannot be stable; but

$$\mathfrak{h}^0=(\mathfrak{e}_{-1}^2\mid X,\operatorname{fil}\{]0,t[\times\{-1,1\}: t>0\})$$

is a non-convergent fully free Cauchy filter pair.

c) O, S, *not* A. On $X=(-1/\mathbb{N}\cup 1/\mathbb{N})\times\{0\}\cup\mathbb{R}\times\{1\}$, let $\mathcal{U}=\mathcal{U}(d)$, where

$$\begin{aligned} d(x,y)=y'-x' \text{ if either } & x'<0<y',\\ \text{or } & x''=y''=1,\ x'<y',\ x'y'=0,\\ \text{or } & x''=1,\ y''=0,\ x'<y'<0,\\ \text{or } & x''=0,\ y''=1,\ 0<x'<y'. \end{aligned}$$

It is enough to consider Cauchy filter pairs $\mathfrak{f}^0$ with $\overline{\mathbb{R}}_i\times\mathbb{R}\cap X\in\mathfrak{f}^i$ $(i=\pm1)$, since the others are convergent. If $\mathfrak{f}^0$ is open then $\mathbb{R}_i\times\{1\}\in\sec\mathfrak{f}^i$, thus (using the Cauchy property) $(0,1)$ is a cluster point of $\mathfrak{f}^0$. If $\mathfrak{f}^0$ is stable then $\mathbb{R}\times\{0\}\notin\sec\mathfrak{f}^i$ and $(0,1)\in\bigcap\mathfrak{f}^i$ $(i=\pm1)$, and so $\mathfrak{f}^0$ converges to $(0,1)$. But the Cauchy filter pair $\mathfrak{e}_0^2\mid X$ has no cluster point.

d) R, S, *not* O. Consider the subspace $X_0=(-1/\mathbb{N}\cup 1/\mathbb{N})\times\{0,1\}\cup$ $\cup\{(0,1)\}$ of the previous example. R and S follow in the same way as O and S in c). $\mathfrak{e}_0^2\mid X_0$ is again a filter pair having no cluster point, and it is open, thus $\mathcal{U}\mid X_0$ is not O-complete.

e) S, *not* R. On $X=\mathbb{R}_0$, let

$$d(x,y)=y-x \quad \text{if} \quad x<0<y.$$

The round Cauchy filter pair $\mathfrak{e}_0\mid X$ has no cluster point. The non-convergent Cauchy filter pairs are finer than $\mathfrak{e}_0\mid X$, so they are not stable.

f_1) F, *not* S *(totally bounded).* $(X,\mathcal{U})$ from Example 11.1.

f_2) F, *not* S *(T_1)*. On X from a), let

$$d(x,y)=y'-x' \quad \text{if} \quad x'<y'.$$

The non-convergent Cauchy filter pairs are the same as in a); they are now stable (but not free).

g_1) A, *not* SF *(totally bounded)*. Let

$$X=[0,1]\times\{0,1\}\cup[2,3]\times\{2,3\},$$

$$d(x,y)=\begin{cases} y'-x'-1 & \text{if } x'<1<y', \\ |y'-x'| & \text{if } x''=y''. \end{cases}$$

$\mathcal{U}^{\mathrm{stp}}$ is the topological sum of four Euclidean intervals, thus $\mathcal{U}$ is A-complete, since, in totally bounded spaces, A-completeness = L-completeness is equivalent to the compactness of $\mathcal{U}^{\mathrm{stp}}$. But $\mathcal{U}$ is not SF-complete, since

$$(\mathrm{fil}\,\{]1-t,1[\times\{0,1\}: 0<t<1\}, \mathrm{fil}\,\{]2,2+t[\times\{2,3\}: 0<t<1\})$$

is a non-convergent fully free (stable) Cauchy filter pair.

g_2) A, *not* SF *(T_1)*. We modify d on X from b):

$$d(x,y)=y'-x' \text{ if either } x'<0\leqq y', \\ \text{or } x'<y',\ x''=y''.$$

Just as in b), the non-convergent Cauchy filter pairs cluster to $(0,-1)$ and $(0,1)$; but $\mathfrak{h}^0$ from b) is a non-convergent fully free Cauchy filter pair, which is now stable, too.

h) O, U, *not* SA. Let $(X_0,\mathcal{U}_0)$ be the space that was denoted by $(X,\mathcal{U})$ in Example 10.3. On $X=X_0\cup\mathbb{R}$, let $\{U(U_0,\varepsilon)\colon U_0\in\mathcal{U}_0, \varepsilon>0\}$ be a base for $\mathcal{U}$, where

$$\begin{aligned} x\,U(U_0,\varepsilon)\,y \ \text{ iff either } & x\,U_0\,y, \\ \text{or } & x\in\mathbb{R}_{-1}, -\varepsilon<x, \quad y\in X_0, \\ \text{or } & x\in X_0, y\in\mathbb{R}_1, \quad y<\varepsilon, \\ \text{or } & x,y\in\mathbb{R}, \quad -\varepsilon<x\leqq 0\leqq y<\varepsilon, \\ \text{or } & x=y. \end{aligned}$$

$\mathcal{U}$ is indeed a quasi-uniformity, since if $V_0\in\mathcal{U}_0$, $V_0^2\subset U_0$ then $U(V_0,\varepsilon)^2\subset \subset U(U_0,\varepsilon)$. Let $\mathfrak{f}^0$ be an open Cauchy filter pair. If $\mathbb{R}_{-i}\in\mathfrak{f}^i$ for $i=-1$ or 1 then $\mathfrak{f}^0$ is clearly convergent; otherwise, the openness implies that $\mathbb{R}_i\in\sec\mathfrak{f}^i$ $(i=\pm1)$, and so $\mathfrak{f}^0$ clusters to 0. Hence $\mathcal{U}$ is O-complete.

Assume now that $\mathfrak{f}^0$ is a free stable Cauchy ultrafilter pair. No free filter is $\mathcal{U}^i\,|\,\mathbb{R}$-stable, thus $X_0\in\mathfrak{f}^i$ $(i=\pm1)$, and so $\mathfrak{f}^0\,|\,X_0$ is a free $\mathcal{U}_0$-stable Cauchy ultrafilter pair; but we saw in Example 10.3 that there do not exist such filter pairs. Hence $\mathcal{U}$ is U-complete, too.

But $\mathcal{U}$ is not SA-complete, since X_0 is a doubly closed non-SA-complete subspace.

i) R, U, *not* SO. Let X_0, $\mathcal{U}_0$ and X be as in h), and

$$\mathcal{B}=\{U(U_0,F): U_0\in\mathcal{U}_0, \emptyset\neq F\subset X_0,\ F \text{ is finite}\},$$

where, with $\varepsilon=1/|F|$,

$$\begin{aligned} x\,U(U_0,F)\,y \ \text{ iff either } & x\,U_0\,y, \\ \text{or } & x\in\mathbb{R}_{-1}, -\varepsilon<x, \quad y\in X_0\setminus F, \\ \text{or } & x\in X_0\setminus F, \quad y\in\mathbb{R}_1, \quad y<\varepsilon, \\ \text{or } & x,y\in\mathbb{R}, \quad -\varepsilon<x\leqq 0\leqq y<\varepsilon, \\ \text{or } & x=y. \end{aligned}$$

If $U_0'\subset U_0$ and $F'\supset F$ then $U(U_0',F')\subset U(U_0,F)$, thus $\mathcal{B}$ is a filter base. Given $U_0\in\mathcal{U}_0$ and a finite $F\subset X_0$, we can choose a $V_0\in\mathcal{U}_0$ such that $V_0^2\subset\subset U_0$ and $V_0^i x=\{x\}$ for $x\in F$ and $i=\pm1$ (because $\mathcal{U}_0$ induces the discrete bitopology). Now $U(V_0,F)^2\subset U(U_0,F)$, thus $\mathcal{B}$ is a base for a quasi-uniformity $\mathcal{U}$.

Let $\mathfrak{f}^0$ be a round Cauchy filter pair. Just as in the preceding example, we can assume that $\mathbb{R}_{-i}\notin\mathfrak{f}^i$; moreover, it is enough to consider free filter pairs. Thus for each $S\in\mathfrak{f}^i$, $S\cap(X_0\cup\mathbb{R}_i)$ is infinite. If $S\cap X_0$ is infinite then for each $U\in\mathcal{U}$,

$$U^i[S]\cap\mathbb{R}_i\neq\emptyset; \tag{1}$$

if $S\cap X_0$ is finite then clearly $S\cap\mathbb{R}_i\neq\emptyset$, i.e. (1) holds in any case; thus, $\mathfrak{f}^i$ being $\mathcal{U}^i$-round, $\mathbb{R}_i\in\sec\mathfrak{f}^i$ $(i=\pm1)$, implying that 0 is a cluster point of $\mathfrak{f}^0$. Hence $\mathcal{U}$ is R-complete. It is also U-complete, for the same reason as the previous example.

Take a non-clustering stable Cauchy filter pair $\mathfrak{f}^0$ in $(X_0,\mathcal{U}_0)$ (see in Example 10.3). As the points of X_0 are isolated in both topologies of $\mathcal{U}$, the stable Cauchy filter pair $\operatorname{Fil}_X\mathfrak{f}^0$ is open, and 0, the only non-isolated point in X, is not a cluster point of it. Hence $\mathcal{U}$ is not SO-complete.

j) U, *not* SR. Example 10.3. (The filter pair showing that $\mathcal{U}$ is not SA-complete was round.) $\mathcal{U}_{\mathrm{so}}\,|\,\mathbb{R}_0$ is a quasi-metrizable example for the weaker statement L $\not\Rightarrow$ SR.

k) O, L, *not* U. Starting from $(X_0,\mathcal{U}_0)=(\mathbb{R}_0,\mathcal{U}_{\mathrm{so}}\,|\,\mathbb{R}_0)$, which is L-complete but not U-complete, add a copy of $\mathbb{R}$ to it and extend $\mathcal{U}_0$ in the same way as in h). $\mathcal{U}$ is L-complete, because $\mathcal{U}^s$ is the discrete uniformity. O-completeness can be proved like in h). Any ultrafilter pair finer than $\operatorname{Fil}_X(\mathfrak{e}_0\,|\,\mathbb{R}_0)$ shows that $\mathcal{U}$ is not U-complete.

l) O, *not* L. On $X=-1/\mathbb{N}\cup\{0\}\cup 1/\mathbb{N}$, let

$$d(x,y)=\begin{cases}|y-x| & \text{if } xy>0,\\ y-x & \text{if } x<y,\ xy=0,\\ \min\{-x,y\} & \text{if } x<0<y.\end{cases}$$

$\mathcal{U}=\mathcal{U}(d)$ is totally bounded. It is not L-complete, since $\mathcal{U}^{\mathrm{stp}}$ is not compact (an infinite discrete topology). But $\mathcal{U}$ is O-complete:

Let $\mathfrak{f}^0$ be a free open Cauchy filter pair. Now $1/\mathbb{N}\in\sec\mathfrak{f}^1$, since otherwise $-1/\mathbb{N}\in\sec\mathfrak{f}^1$, and then the openness implies again that $1/\mathbb{N}\in\sec\mathfrak{f}^1$ (cf. the last line in the definition of d). Similarly, $-1/\mathbb{N}\in\sec\mathfrak{f}^{-1}$. The free filter base pair $(\mathfrak{f}^{-1}\mid -1/\mathbb{N},\ \mathfrak{f}^1\mid 1/\mathbb{N})$ converges to 0, so $\mathfrak{f}^0$ clusters to it. Hence $\mathcal{U}$ is O-complete.

m) R, *not* LO. On $X=[-1,1]\times\{0\}\cup\{0\}\times]0,1]$, let

$$d(x,y)=d^2_{\mathrm{eu}}(x,y)\ \text{ if either } x'y'>0,\\ \text{or } x'',y''>0,\\ \text{or } x'<0\leqq y',\\ \text{or } x'=0<y'.$$

$\mathfrak{f}^0=\mathrm{Fil}\ (\mathfrak{e}_0\times\mathfrak{e}_1\mid\{0\}\times]0,1])$ is an open linked Cauchy filter pair having no cluster point. But $\mathcal{U}$ is R-complete, since 0 is a cluster point of any non-convergent round Cauchy filter pair.

n) R, LO, *not* O *(totally bounded)*. With X_0 the set and d_0 the distance that were denoted in l) by X and d, define on $X=X_0\times\{0,1\}$:

$$d(x,y)=\begin{cases}d_0(x',y') & \text{if either } x''=y'',\\ & \quad\text{or } x'<0<y',\ x''\neq y'',\\ x'+y' & \text{if } x'\geqq 0<y',\ x''=1,\ y''=0.\end{cases}$$

It is straightforward to check the Triangle Inequality. Let $\mathfrak{f}^0$ be an open linked Cauchy filter pair. Then for $k=0$ or 1, $X_0\times\{k\}\in\sec\mathfrak{f}^{-1}(\cap)\mathfrak{f}^1$, thus $\mathfrak{f}^0\mid X_0\times\{k\}$ is an open Cauchy filter pair in a copy of the O-complete space from l). Hence $\mathfrak{f}^0$ has a cluster point, and $\mathcal{U}$ is LO-complete.

Let now $\mathfrak{f}^0$ be a free round Cauchy filter pair. If $\mathfrak{f}^0\mid X_0\times\{k\}$ is a filter pair for $k=0$ or 1 then it is a round, hence open, Cauchy filter pair in a copy of the O-complete space from l), therefore it has a cluster point. If $X_0\times\{0\}\in\sec\mathfrak{f}^{-1}$ and $X_0\times\{1\}\in\sec\mathfrak{f}^1$ then the $\mathcal{U}$-roundness of $\mathfrak{f}^1$ implies that $X_0\times\{0\}\in\sec\mathfrak{f}^1$, too; thus $\mathfrak{f}^0\mid X_0\times\{0\}$ is a filter pair, and this case has already been dealt with. Finally, assume that $X_0\times\{1\}\in\sec\mathfrak{f}^{-1}$, $X_0\times\times\{0\}\in\sec\mathfrak{f}^1$. As $\mathfrak{f}^{-1}$ is $\mathcal{U}^{-\mathrm{tp}}$-open, we have $-1/\mathbb{N}\times\{1\}\in\sec\mathfrak{f}^{-1}$; the Cauchy property implies that $1/\mathbb{N}\times\{0\}\in\sec\mathfrak{f}^1$. Now $(0,1)$ is a cluster point of $\mathfrak{f}^0$. Hence $\mathcal{U}$ is R-complete.

But $\mathcal{U}$ is not O-complete: the open Cauchy filter pair generated by the sequences $-1/\mathbb{N}\times\{0\}$ and $1/\mathbb{N}\times\{1\}$ has no cluster point.

o) LO, *not* R *(totally bounded).* With X_0, d_0 and X from the previous example, let

$$d(x,y)=\begin{cases} d_0(x',y') & \text{if } x''=y''=0,\\ y'-x' & \text{if } x'<0<y',\ x''=0,\ y''=1.\end{cases}$$

LO-completeness follows in the same way as in n). The filter pair considered at the end of n) is now round, but it has no cluster point. □

14.3 PROPOSITION. *U-complete quasi-pseudometric spaces are* SR*-complete.*

PROOF. Let $\mathfrak{f}^0$ be a round stable Cauchy filter pair. Define

$$A_n=\bigcap\{U_{(2^{-n})}[S]\colon S\in\mathfrak{f}^1\}\qquad(n\in\mathbb{N}).$$

$A_n\in\mathfrak{f}^1$, so $U_{(2^{-n})}[A_{n+1}]\supset A_n$, and we can pick points $x_n\in A_n$ (starting with an arbitrary $x_1\in A_1$) such that $d(x_{n+1},x_n)<2^{-n}$. By the Triangle Inequality, $d(x_k,x_n)<2^{-n+1}$ for $k\geq n$. Given an $F\in\mathfrak{f}^1$, there are $S\in\mathfrak{f}^1$ and $n\in\mathbb{N}$ such that $U_{(2^{-n})}[S]\subset F$, therefore $A_k\subset F$, and so $x_k\in F$, for each $k\geq n$. Hence the filter $\mathfrak{g}^1$ generated by the sequence $\langle x_n\rangle$ is finer than $\mathfrak{f}^1$. Take an ultrafilter $\mathfrak{h}^1$ finer than $\mathfrak{g}^1$. The filter $\mathfrak{h}^1$ is stable, since the sequence $\langle x_n\rangle$ is frequently in any $S\in\mathfrak{h}^1$, and so

$$\bigcap\{U_{(2^{-n})}[S]\colon S\in\mathfrak{h}^1\}\supset\{x_{n+1},x_{n+2},\dots\}\in\mathfrak{g}^1\subset\mathfrak{h}^1.$$

Similarly, there is a $\mathcal{U}^{-1}$-stable ultrafilter $\mathfrak{h}^{-1}$ finer than $\mathfrak{f}^{-1}$. Now $\mathfrak{h}^0$ is a stable Cauchy ultrafilter pair finer than $\mathfrak{f}^0$; so $\mathfrak{f}^0$ has a cluster point, since $\mathfrak{h}^0$ is convergent. □

PROBLEM. Is each U-complete quasi-pseudometric space SA-complete (or at least SO-complete)?

§ 15 Complete quasi-proximities

15.1 It is natural to call a quasi-proximity δ C-*complete,* F-*complete*, etc. if the totally bounded quasi-uniformity inducing δ has this property. As mentioned in 14.1, $\mathsf{C}=\mathsf{S}$, $\mathsf{F}=\mathsf{SF}$ and $\mathsf{A}=\mathsf{SA}=\mathsf{U}=\mathsf{L}$ for quasi-proximities. It is well-known that δ is L-complete iff δ^{sp} is compact; if $\mathcal{U}$ is totally bounded then so is ${}^{\mathsf{L}}\mathcal{U}$, and so each quasi-proximity has an L-completion with fairly good properties. (For details, see e.g. [Cs2] § 16.) It is the aim of this section to show that F- or C-completions cannot be as good as that. Total boundedness is preserved by the construction ${}^{\mathsf{S}}\mathcal{U}$ (Proposition 6.5), so each quasi-proximity has a C-completion, which is, however, not even weakly basic. On the other hand, ${}^{\mathsf{SF}}\mathcal{U}$ does not preserve total boundedness, while its totally bounded reflexion is not necessarily (S)F-complete; the proof of Theorem 15.2 will contain a totally bounded $\mathcal{U}$ with these properties.

Let us also remark that the completeness properties of quasi-proximities could also be defined without using quasi-uniformities. A filter pair $\mathfrak{f}^0$ in (X,δ) is called *compressed* ([De3] 5.1) if $A \in \sec \mathfrak{f}^{-1}$, $B \in \sec \mathfrak{f}^1$ imply $A\,\delta\,B$. According to [De3] Lemma 5.1, if $\mathcal{U}$ is totally bounded then $\mathcal{U}$-Cauchy $=\mathcal{U}^{\mathrm{t}}$-compressed. Moreover, the roundness of filter pairs in a quasi-uniform space is a quasi-proximity invariant ($\mathfrak{f}^i$ is $\mathcal{U}^i$-round iff for any $S \in \mathfrak{f}^i$ there is a $T \in \mathfrak{f}^i$ such that $T\,\overline{\mathcal{U}^{i\mathrm{t}}}\,X \setminus S$). Hence we can say in terms of quasi-proximities that δ is C-complete (F-complete) iff each compressed (fully free compressed) filter pair is convergent; R-complete iff each round compressed filter pair has a cluster point; etc.

15.2 THEOREM. *There is no natural basic bitopological* C*-completion or* F*-completion for quasi-proximities.*

PROOF. It is enough to prove the theorem for F-completeness, since a C-completion would also be an F-completion (and the notions natural, basic and bitopological are independent of the completeness in question). If there existed a natural basic bitopological F-completion then an extension $({}^{\mathsf{F}}X, {}^{\mathsf{F}}\mathcal{U})$ could be assigned to each totally bounded space $(X,\mathcal{U})$ such that

(i) ${}^{\mathsf{F}}\mathcal{U}$ is totally bounded;

(ii) the trace filter pairs of the new points are not convergent;

(iii) there belong different trace filter pairs to different new points;

(iv) any $(\mathcal{U},\mathcal{V})$-isomorphism can be extended to an $({}^{\mathsf{F}}\mathcal{U}, {}^{\mathsf{F}}\mathcal{V})$-isomorphism;

(v) the identity of X can be extended to an isomorphism between $({}^{\mathsf{F}}X, {}^{\mathsf{F}}\mathcal{U}^{-1})$ and $({}^{\mathsf{F}}X, {}^{\mathsf{F}}(\mathcal{U}^{-1}))$ (although not shown in the notation, ${}^{\mathsf{F}}X$ does not mean the same in the two extensions).

With the notations of Example 6.6, let $v,w \notin \mathbb{R}^3$, $v \neq w$, $X = X^{\circ\circ} \times 1/\mathbb{N} \cup \cup \{v,w\}$, and

$$d(x,y) = \begin{cases} d^{\circ} \times d_{\mathrm{eu}}(x,y) & \text{if } x,y \in \mathbb{R}^3, \\ d^3_{\mathrm{eu}}(x,0^3) & \text{if } x \in \mathbb{R}^3,\ x' < 0,\ y \notin \mathbb{R}^3, \\ d^3_{\mathrm{eu}}(0^3,y) & \text{if } x \notin \mathbb{R}^3,\ y \in \mathbb{R}^3, y' > 0, \\ 0 & \text{if } x = v,\ y = w. \end{cases}$$

([De7] Example 3.2 is very similar, but $d(x,w)$ and $d(v,y)$ are not the same as here.) $\mathcal{U} = \mathcal{U}(d)$ is totally bounded. Define

$$\mathfrak{h}^0_n = \mathrm{Fil}\ (\mathfrak{e}_0 \times \mathfrak{e}^2(0^2, 1/n) \mid \mathbb{R}^3 \cap X) \qquad (n \in \mathbb{N}).$$

These filter pairs are non-convergent, fully free and minimal Cauchy. For $n \in \mathbb{N}$, let p_n denote a point in ${}^{\mathsf{F}}X \setminus X$ to which $\mathfrak{h}_n^0$ converges (there are such points, since ${}^{\mathsf{F}}\mathcal{U}$ was assumed to be F-complete and $\mathfrak{h}_n^0$ is fully free in ${}^{\mathsf{F}}X$, cf. the proof of Lemma 11.2). The trace filter pair $\mathfrak{f}^0(p_n)$ is Cauchy, so the minimality of $\mathfrak{h}_n^0$ implies that $\mathfrak{f}^0(p_n) = \mathfrak{h}_n^0$. (There have to be more points in ${}^{\mathsf{F}}X \setminus X$, but we shall not need them.)

To avoid ambiguity, let us denote the fundamental set of ${}^{\mathsf{F}}(\mathcal{U}^{-1})$ by Y, and the trace filter pairs belonging to this extension by $\mathfrak{g}^0(a)$ $(a \in Y)$. The filter pairs $\mathfrak{k}_n^0 = (\mathfrak{h}_n^1, \mathfrak{h}_n^{-1})$ $(n \in \mathbb{N})$ are non-convergent, fully free and minimal Cauchy in $(X, \mathcal{U}^{-1})$, thus $\mathfrak{g}^0(q_n) = \mathfrak{k}_n^0$ with suitable points $q_n \in Y \setminus X$.

Let now $\mathfrak{f}$ be a free ultrafilter in ${}^{\mathsf{F}}X$ such that $P = \{p_n : n \in \mathbb{N}\} \in \mathfrak{f}$. By (i), ${}^{\mathsf{F}}\mathcal{U}$ is totally bounded, so $(\mathfrak{f}, \mathfrak{f})$ is a Cauchy filter pair. $\mathfrak{f}$ is also fully free, since ${}^{\mathsf{F}}\mathcal{U} \mid P$ is T_0 (otherwise, two different trace filters $\mathfrak{f}^1(p_n)$ and $\mathfrak{f}^1(p_m)$ would be comparable). As ${}^{\mathsf{F}}\mathcal{U}$ is F-complete, $(\mathfrak{f}, \mathfrak{f})$ converges to some $a \in$ $\in {}^{\mathsf{F}}X$. Let $\mathfrak{f}^0$ be the ${}^{\mathsf{F}}\mathcal{U}$-envelope of $(\mathfrak{f}, \mathfrak{f})$; then $\mathfrak{f}^0$ also converges to a. Given $\varepsilon > 0$, take $K = K(\varepsilon) \in \mathfrak{f}^{\mathsf{X}}(a)$ such that $K \subset U_{(\varepsilon)}$. Each $S \in \mathfrak{f}^1$ contains an ${}^{\mathsf{F}}\mathcal{U}^{\mathrm{tp}}$-neighbourhood of an infinite subset of P, so $S \cap X \in \mathfrak{f}^1(p_n)$ for infinitely many $n \in \mathbb{N}$. As $\mathfrak{f}^1$ ${}^{\mathsf{F}}\mathcal{U}^{\mathrm{tp}}$-converges to a, $K_1 = S \cap X$ with a suitable $S \in \mathfrak{f}^1$; so there are $y_1, y_2 \in K_1$ such that $y_1', y_2' > 0$, $y_1'' < 0 < y_2''$, $y_1''', y_2''' < \varepsilon$. Similarly, there are $x_1, x_2 \in K_{-1}$ with $x_1', x_2' < 0$, $x_1'' < 0 < x_2''$, $x_1''', x_2''' < \varepsilon$. Together with $K \subset U_{(\varepsilon)}$, this means that

$$K_i \subset (]0, i\varepsilon[\times] - \varepsilon, \varepsilon[\times]0, 2\varepsilon[\cap X) \cup \{v, w\} \qquad (i = \pm 1).$$

Moreover, if $w \in K_{-1}$ then $v \notin K_1$. As there is such a $K(\varepsilon)$ for each $\varepsilon > 0$, we obtain that $\mathfrak{f}^0(a)$ converges to v or w; thus $a \in X$ by (ii). $(\mathfrak{f}, \mathfrak{f})$ clearly does not converge to any point in $\mathbb{R}^3$, so $a = v$ or $a = w$. Assume for a moment that $(\mathfrak{f}, \mathfrak{f})$ converges to both points, and take $V \in {}^{\mathsf{F}}\mathcal{U}$ such that $V^2 \mid X \subset U_{(1)}$. As $\mathfrak{f}$ ${}^{\mathsf{F}}\mathcal{U}^{\mathrm{tp}}$-converges to w and ${}^{\mathsf{F}}\mathcal{U}^{-\mathrm{tp}}$-converges to v, we have $Vw \cap V^{-1}v \in \mathfrak{f}$, so this set is non-empty, implying that $w\, U_{(1)}\, v$, a contradiction. Hence $(\mathfrak{f}, \mathfrak{f})$ converges to exactly one point in ${}^{\mathsf{F}}X$, which is either v or w.

Let $Q = \{q_n : n \in \mathbb{N}\}$, and denote by $\mathfrak{g}$ the ultrafilter in Y for which $\mathfrak{g} \mid Q$ is the image of $\mathfrak{f} \mid P$ under the bijection $p_n \mapsto q_n$ $(n \in \mathbb{N})$. Just as above, $(\mathfrak{g}, \mathfrak{g})$ converges to exactly one point, which is either v or w, say it is v. (The other case can be similarly dealt with.) Let $f : Y \to {}^{\mathsf{F}}X$ be an $({}^{\mathsf{F}}(\mathcal{U}^{-1}), {}^{\mathsf{F}}\mathcal{U}^{-1})$-isomorphism according to (v). For $n \in \mathbb{N}$ fixed, let $a = f(q_n)$. The ${}^{\mathsf{F}}\mathcal{U}^{-1}$-trace filter pair of a is $(\mathfrak{f}^1(a), \mathfrak{f}^{-1}(a))$, which has to coincide with $\mathfrak{g}^0(q_n)$ (since f is an isomorphism), so $\mathfrak{f}^0(a) = \mathfrak{f}^0(p_n)$ (see the definition of $\mathfrak{g}^0(q_n)$), implying $a = p_n$ by (iii). Thus $f(q_n) = p_n$ $(n \in \mathbb{N})$, and so $f(\mathfrak{g}) = \mathfrak{f}$. Consequently, $(\mathfrak{f}, \mathfrak{f})$ converges in $({}^{\mathsf{F}}X, {}^{\mathsf{F}}\mathcal{U}^{-1})$, so also in $({}^{\mathsf{F}}X, {}^{\mathsf{F}}\mathcal{U})$, to $f(v) = v$.

On the other hand, the $(\mathcal{U}^{-1},\mathcal{U})$-isomorphism h defined by

$$h(x)=\begin{cases}(-x',x'',x''') & \text{if } x\in\mathbb{R}^3,\\ w & \text{if } x=v,\\ v & \text{if } x=w,\end{cases}$$

has, according to (iv), an $({}^{\mathsf{F}}(\mathcal{U}^{-1}),{}^{\mathsf{F}}\mathcal{U})$-isomorphic extension g, and an argument similar to the one in the preceding paragraph yields now that $(\mathfrak{f},\mathfrak{f})$ converges to $g(v)=w$, a contradiction. □

There is still a gap between this theorem and the positive statement that a C-completion ${}^{\mathsf{C}}\delta$ can be obtained through ${}^{\mathsf{S}}\mathcal{U}$, since ${}^{\mathsf{C}}\delta$ is not even weakly basic.

15.3 A quasi-proximity δ is L-complete iff the proximity δ^{s} is complete, i.e. iff δ^{sp} is compact; this topology is the supremum of the topologies $\delta^{-\mathrm{p}}$ and δ^{p}, so δ is L-complete iff the induced bitopology is sup-compact. Assume now that δ is C-complete (or F-complete). Then it is L-complete, hence its bitopology $(\mathcal{T}^{-1},\mathcal{T}^1)=(\delta^{-\mathrm{p}},\delta^{\mathrm{p}})$ is sup-compact, which implies that δ is the only quasi-proximity compatible with $(\mathcal{T}^{-1},\mathcal{T}^1)$ (a condition weaker than sup-compactness is sufficient here, see [JB] Theorem 2.4). So we might call a completely regular (=quasi-uniformizable) bitopology C-*compact*, respectively F-*compact* (properties stronger than sup-compactness) if it has a (unique) compatible C-complete, respectively F-complete, quasi-proximity. The following question is now quite natural: can the C-compact (F-compact) bitopological spaces be characterized without using quasi-uniformities (e.g. in terms of covers)? C-compactness seems to be very problematic, since the finite bitopological space in Example 11.1 is not C-compact.

§ 16 Summing up

Similarly to the table in 2.2, we summarize the properties of the bitopological notions of completeness and of the corresponding completions. a), b), f) and g) from § 1 are satisfied in each case. All the completions are natural and, excepting ${}^{\mathsf{S}}\mathcal{U}$, basic. (Notation: $\mathfrak{P}_{\mathsf{SO}}=\mathfrak{P}_{\mathsf{S}}\cap\mathfrak{P}_{\mathsf{O}}$, $\mathfrak{P}_{\mathsf{LO}}=\mathfrak{P}_{\mathsf{L}}\cap\mathfrak{P}_{\mathsf{O}}$.)

completeness	c)	d)	e)	completion	h)	i)
C	+	+	+	?		
F	+	+	+	?		
A	+	+	+	?		
O	−	+	−	${}^{\mathsf{O}}\mathcal{U}={}^{\mathsf{O}}\mathcal{U}(\mathfrak{P}_{\mathsf{O}}^{\mathrm{MnE}})$	−	+
R	−	+	−	${}^{\mathsf{R}}\mathcal{U}$	+	+
S	+	+	+	${}^{\mathsf{S}}\mathcal{U}$	−	−
SF	+	+	+	${}^{\mathsf{SF}}\mathcal{U}$	+	+

SA	+	+	+	${}^{\mathsf{SA}}\mathcal{U}$	−	−
SO	−	+	−	${}^{\mathsf{SO}}\mathcal{U} = {}^1\mathcal{U}(\mathfrak{P}^{\mathrm{NE}}_{\mathsf{SO}})$	−	−
SR	−	+	−	${}^{\mathsf{SR}}\mathcal{U}$	−	−
U	+	+	+	${}^{\mathsf{U}}\mathcal{U}$	+	+
L	+	−	+	${}^{\mathsf{L}}\mathcal{U}$	+	+
LO	−	−	−	${}^{\mathsf{LO}}\mathcal{U} = {}^4\mathcal{U}(\mathfrak{P}^{\mathrm{NE}}_{\mathsf{LO}})$	−	−
LR	−	−	−	${}^{\mathsf{LR}}\mathcal{U} = {}^4\mathcal{U}(\mathfrak{P}^{\mathrm{EN}}_{\mathsf{L}})$	−	−

To O: Adopt the reasoning used in § 13 for R-completeness. ${}^{\mathsf{O}}\mathcal{U}$ is not an O-complete hull, because two elements of $\mathfrak{P}^{\mathrm{Mn}}_{\mathsf{O}}$ can have different but comparable envelopes, making one of the new points superfluous: let

$$X = \mathbb{R}_0 \times \{0\} \cup \{(t,t) : t \in \mathbb{R}_1\}, \ \mathcal{U} = \mathcal{U}(d^2_{\mathrm{so}} \mid X),$$

$\mathfrak{f}^0, \mathfrak{g}^0 \in \mathfrak{P}^{\mathrm{M}}_{\mathsf{O}}$ such that $\mathfrak{f}^{-1}, \mathfrak{g}^{-1} \supset \mathfrak{e}^2 \mid \mathbb{R}_{-1} \times \{0\}$, $\mathfrak{f}^1 \supset \mathfrak{e}^2 \mid \mathbb{R}_1 \times \{0\}$, $\mathfrak{g}^1 \supset \mathfrak{e}^2 \mid \mid \{(t,t) : t \in \mathbb{R}_1\}$.

To SO: The situation is analogous to the case of SR-completeness, cf. 13.3.

To LO *and* LR: The trace filter pairs being linked, ${}^4\mathcal{U}$ can of course be replaced by several other constructions. ${}^{\mathsf{LO}}\mathcal{U}$ and ${}^{\mathsf{LR}}\mathcal{U}$ are neither finest nor complete hulls:

EXAMPLE. On $X = \mathbb{R}_0 \times \{0\} \cup \{0\} \times (-1/\mathbb{N} \cup 1/\mathbb{N})$, let

$$d(x,y) = \begin{cases} d_{\mathrm{eu}}(x'', y'') & \text{if } x''y'' > 0, \\ d_{\mathrm{so}}(x', y') & \text{if } x''y'' = 0. \end{cases}$$

$\mathfrak{P}^{\mathrm{EN}}_{\mathsf{L}} = \mathfrak{P}^{\mathrm{N}}_{\mathsf{LO}} = \mathfrak{P}^{\mathrm{NE}}_{\mathsf{LO}}$ consists of the following two filter pairs: $\mathfrak{e}_0 \times \mathfrak{e}_i \mid X$ $(i = \pm 1)$. It is enough to take only one of the new points to make the space LR-complete (LO-complete). ${}^{\mathsf{LO}}\mathcal{U} = {}^{\mathsf{LR}}\mathcal{U}$ is not finest, since the identity of X cannot be extended to a uniformly continuous map from ${}^{\mathsf{LO}}X$ into the space mentioned in the preceding sentence. □

(Received November 12, 1991)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H-1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 405–413

THE CROSSING NUMBER OF A CUBIC PLANE POLYHEDRAL MAP PLUS AN EDGE

A. RISKIN

1. Introduction

It follows from a classical result of Whitney [8] that planar 3-connected graphs have unique plane embeddings (where *unique* means that in any two embeddings, the regions are bounded by the same cycles). Clearly then joining any two vertices which do not lie in the same region of this unique embedding yields a nonplanar graph. In this paper, we calculate the crossing numbers of all graphs obtained in this way from *cubic* 3-connected planar graphs. These crossing numbers turn out to be exactly what they ought to be; that is, the minimum number of crossings is the same as the minimum number of crossings introduced by adding the edge without redrawing the unique embedding of the original cubic graph. We conjecture that the result still holds for noncubic plane polyhedral maps. However, it is unfortunate that the generalization to the case of adding more than one edge is false, as the example in Figure 1 shows. This is rather disappointing in that that (false)

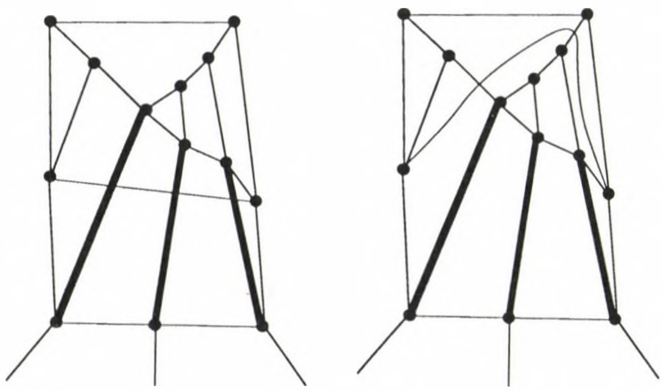

Fig. 1. Three added edges (bold) but crossing number $\leqq 2$

1991 *Mathematics Subject Classification.* Primary 05C10.
Key words and phrases. Crossing number, planar, 3-connected, polyhedral.

generalization would have sufficed to settle a number of unsolved problems in the area of crossing numbers; most prominently the twenty year old problem proposed by Harary et al. [4] of showing that $\mathrm{cr}(C_m \times C_n) = n(m-2)$ for $m \leqq n$. Note that this was solved for $m = 3, 4$ by Beineke and Ringeisen [2, 5].

2. Definitions and preliminary results

In this paper, graphs have no loops, multiple edges, or 2-valent vertices. All standard graph theory terms conform to [3] unless otherwise mentioned. A *map* is an embedding of a graph G in a surface M. The *faces* of the map are the connected components of $M - G$. Note that when the meaning is clear from the context, we confuse a face with its bounding circuit, or with its closure, as in the next definition. A map is *polyhedral* if no two faces have a multiply connected union. A *cellular subcomplex* of a map is a set of faces whose union is homeomorphic to a 2-cell. *Removing* an edge of a graph means doing just that, along with coalescing any resulting 2-valent vertices into the edges in which they lie. *Splitting a face* of a map is the operation inverse to removing an edge of a map. A *cubic* face splitting is one in which two new 3-valent vertices are created.

An edge of a polyhedral map is *removable* if removing it yields a polyhedral map. Two faces of a map are said to meet *improperly* if they have a multiply connected union. An *obstacle* to the removal of an edge e of a polyhedral map is a face which meets improperly the new face created upon the removal of edge e. A *3-chain* is a set of three faces of a map, each of which intersects the other two. If the three faces have a vertex in common, the 3-chain is said to be *trivial*. An *n-chain* is a set of n faces $F_1, \dots, F_n$ such that $F_i \cap F_{i+1} \neq \emptyset$ for $i = 1, \dots, n-1$, and $F_n \cap F_1 \neq \emptyset$. If $\bigcup_{i=1}^{n} F_i$ is a cellular subcomplex then the n-chain is said to be *trivial*. It is a result of the author's [6] that a face F of a map $M \neq K_4$ constitutes an obstacle to the removal of an edge e if and only if F lies in a nontrivial 3-chain with the two faces containing e.

It is a theorem of Steinitz [7], [1] that any planar 3-connected graph is the *1-skeleton* of a 3-polytope (i.e. the graph induced by the vertices and edges of the polytope). Since it is clear that the 1-skeleton of a 3-polytope is a polyhedral map on the sphere, it follows that the unique plane embedding of a planar 3-connected graph must be polyhedral. It is trivial to show that a plane polyhedral map is 3-connected, and thus the two concepts are equivalent for the plane. We therefore tend to use them interchangeably, depending on which conception is more convenient. It is another theorem of Steinitz [7], [1] that all plane polyhedral maps can be generated from K_4 by face splittings. We use this theorem as a means of induction.

Now, let u and v be vertices of a planar 3-connected graph G. We define the *limbic distance* between u and v, denoted by $\lambda(u,v)$ or $\lambda_G(u,v)$, to be the minimum value of n such that there is a chain of faces $F_1, F_2, \ldots, F_n$ of the unique embedding such that $u \in F_1$, $v \in F_n$, and $F_i \cap F_{i+1} \in E(G)$ for $i = 1, \ldots, n-1$. Such a minimum length chain is called a *limbic face chain* joining u and v, and an edge is called *limbic* with respect to u and v if it is the intersection of two faces in a limbic face chain joining u and v.

Note that the same definition will serve for *limbic distance* in a map M, denoted by $\lambda_M(u,v)$, although this value will be dependent on the embedding, rather than on the graph. A *drawing* of a graph G is an immersion of G in a surface which is no more than 2 to 1, and avoids the pathology of vertices being mapped on to edges. In this paper, all drawings will be on the plane (or, equivalently, the sphere). The *crossing number* of a drawing D, denoted by $\mathrm{cr}(D)$, is the number of edge crossings in the drawing. The *crossing number* of a graph G, denoted by $\mathrm{cr}(G)$, is the minimum value of $\mathrm{cr}(D)$ over all drawings of G.

The *derived map* of a drawing is the map obtained by adding a vertex at each edge crossing. The *limbic distance* between vertices in a drawing D, denoted by $\lambda_D(u,v)$, is the limbic distance between them in the derived map. Note that the number of edge crossings introduced in joining two nonadjacent vertices u and v in a drawing D is $\lambda_D(u,v) - 1$.

3. Classification of noninductive graphs

Throughout the rest of the paper, all graphs are cubic, planar, and 3-connected and all face splittings are cubic. A *triangulogenic* cubic face splitting is one in which the new edge is added across a face between two incident edges. A *nontriangulogenic* cubic face splitting is one in which the two edges joined are not incident. The three edges *involved* in a triangulogenic face splitting are the two incident edges joined by the new edge, along with the third edge of the original map incident to both of them. The two edges *involved* in a nontriangulogenic face splitting are the two nonincident edges joined.

Lemma 1. *If one of the two edges involved in a nontriangulogenic splitting is removable, then it is split into two removable edges, and any removable edge not involved in the splitting is removable after the splitting.*

Proof. If faces F, G and H lie in a nontrivial 3-chain, and removing a removable edge e destroys it, then clearly $F \cup G \cup H$ must surround a triangular face which contains edge e on its bounding circuit. Thus undoing a nontriangulogenic face splitting cannot destroy any nontrivial 3-chains, and so if an edge or portion of an edge is nonremovable after a nontriangulogenic splitting, it must have been nonremovable before it. □

Lemma 2. *Any removable edge not involved in a triangulogenic splitting is removable after the splitting.*

PROOF. Similar to the proof of Lemma 1. □

LEMMA 3. *In a cubic polyhedral map other than K_4 all edges lying in triangular faces are removable.*

PROOF. This follows directly from a result of the author's [6] which states that if a triangular face of a polyhedral map has a 3-valent vertex v in it, the edge opposite v is removable. □

COROLLARY 1. *After a triangulogenic splitting, the three edges of the generated triangle are removable, whereas the three edges incident to but not contained in the generated triangle are not removable.*

THEOREM 1. *Every cycle in a cubic plane polyhedral map other than K_4 contains a removable edge.*

PROOF. We proceed by induction on the number of edges, tacitly invoking the theorem of Steinitz [7], [1] mentioned above. If there are 7 edges, we have the 1-skeleton of the triangular prism, for which the theorem is clearly true. Now suppose that the theorem is true for cubic plane polyhedral maps with $n-1$ edges, and let M be such a map with $n>7$ edges. Further suppose that M is obtained from a map N by a face splitting consisting of adding an edge e.

Now, if adding e to N consists in a nontriangulogenic face splitting, then any edge of M which was removable in N is still removable in M. Thus any cycle of M which does not contain edge e is a cycle in N, and so by the inductive hypothesis has a removable edge, which is still removable in M. Also, any cycle in M which does contain edge e has a removable edge, namely e.

On the other hand, if adding edge e to N consists in a triangulogenic splitting, then any cycle of M which does not contain any of the 6 edges in or incident to the generated triangle is a cycle in N, and thus contains a removable edge by the inductive hypothesis. Any cycle which does contain one of these edges must contain one of the three edges in the generated triangle, which are all removable by Lemma 3. □

Note that Theorem 1 is no longer true if the word "cubic" is removed from the hypotheses, as the graph of the bipyramid over a triangle shows.

A cubic planar 3-connected graph G is called *noninductive* with respect to vertices u and v in case all removable edges not incident with u or v are limbic with respect to u and v. We now proceed to classify all such graphs. The limbic distance between a vertex v and a face F, denoted by $\lambda(v,F)$, is the limbic distance between v and a new vertex in the interior of F. The k^{th} *u-ripple*, denoted by R_u^k, is $\{e \in E(G) \mid \lambda(u,F)=k$ and $\lambda(u,G)=k+1$ where $F \cap G = e\}$. An *index k u-strut* is an edge e such that $e \cap R_u^k \neq \emptyset$ and $e \cap R_u^k \neq \emptyset$.

LEMMA 4. *In a cubic planar 3-connected graph G, the k^{th} u-ripple is a (possibly empty) union of disjoint cycles.*

PROOF. If $R_u^k = \emptyset$, we are done, thus consider an edge $e \in R_u^k$. Let the endpoints of e be vertices x and y, and let the three faces containing x be F, G, and H, where $F \cap G = e$. Furthermore, let $F \cap H = f$ and $G \cap H = g$. We will show that exactly one of the two edges at each endpoint of e is also in R_u^k. Suppose $\lambda(u, F) = k$ and $\lambda(u, G) = k + 1$. Clearly $k \leqq \lambda(u, H) \leqq k + 1$. In the first possibility, $g \in R_u^k$ whereas in the second, $f \in R_u^k$. The same argument works for the other two edges incident with vertex y. □

LEMMA 5. *The different components of R_u^k, if any, do not lie within one another.*

PROOF. Clear. □

LEMMA 6. *If G is noninductive with respect to u and v, with $\lambda(u, v) = n$, then R_u^k for $1 \leqq k \leqq n - 1$ consists of a single cycle which separates u from v, and $R_u^k = \emptyset$ for $k \geqq n$.*

PROOF. Let $1 \leqq k \leqq n - 1$. If R_u^k contains a component C which does not separate u and v, then none of the edges of C are limbic with respect to u and v. By Theorem 1, C has a removable edge, contradicting the assumption that G is noninductive. The other assertions in the lemma are trivial. □

On the strength of the foregoing lemmas, we make the following definition: If G is noninductive with respect to u and v, then k^{th} *annulus*, denoted by A_k, for $k = 1, 2, \dots, \lambda(u, v) - 2$, is the annulus bounded by R_u^k and R_u^{k+1}. We define A_0 to be the star of u and $A_{\lambda(u,v)-1}$ to be the star of v.

LEMMA 7. *Let $P \subset A_k$ be a path with initial vertex in R_u^k and terminal vertex in R_u^{k+1}. Then P consists of a single edge.*

PROOF. Suppose not. Then there is a vertex $w \in P$ lying in the interior of A_k. Let e_1 and e_2 be the two edges of P incident with w. Then each of the e_i's determines a subpath P_i running from w to the boundary of A_k. We assume that P_1 runs from w to R_u^k whereas P_2 runs from w to R_u^{k+1}. Now let e_3 be the third edge incident with w. We claim that e_3 determines a third path P_3 from w to the boundary of A_k, and also that $P_3 \cap P_i = \{w\}$ for $i = 1, 2$.

First, if every path Q beginning with e_3 has empty intersection with the boundary of A_k, then either Q separates R_u^k from R_u^{k+1} or else Q bounds a cellular region in the interior of A_k. Clearly neither of these is possible. Thus there is a path containing e_3 running from w to the boundary of A_k. Let P_3 be the shortest such path. If $\{w\} \subsetneqq P_3 \cap P_i$ for $i = 1$ or 2, then again, either P_3 along with a segment of the interior of P separates R_u^k from R_u^{k+1} or else there is a cellular subcomplex contained in the interior of A_k, neither of which is possible.

Next we claim that $P_3 \cap \partial A_k \subseteqq R_u^k$. If not, then $P_3 \cap \partial A_k \subseteqq R_u^{k+1}$, in which case $P_2 \cup P_3$ along with an appropriate segment of R_u^{k+1} bounds a cellular subcomplex D of G contained in A_k with $D \cap R_u^k = \emptyset$. Let $F \subseteqq D$ be a face of G which intersects R_u^{k+1} in an edge. Thus $\lambda(u, F) = k$, contradicting the assumption that $F \cap R_u^k = \emptyset$. Thus $P_1 \cup P_3$ along with an appropriate segment of R_u^k bounds a cellular subcomplex E of G with $E \subseteqq A_k$ and $E \cap R_u^{k+1} = \emptyset$. Clearly no edge in the boundary of E is limbic with respect to u and v. However, one of those edges is removable by Theorem 1, contradicting the assumption of noninductivity. □

COROLLARY 2. *The only paths across A_k are index k u-struts.*

LEMMA 8. *There are exactly three index k u-struts across each A_k.*

PROOF. If there are < 3, then two faces of G meet improperly in A_k. Thus there are $\geqq 3$. If there are $n \geqq 4$, then due to Corollary 2, A_k consists of a nontrivial n-chain. Since none of the index k u-struts are limbic with respect to u and v, they must all be nonremovable. Let e be one of the index k u-struts, and let F be an obstacle to its removal. Clearly $F \not\subset A_k$, so without loss of generality, assume $F \subset A_{k+1}$. Because F and the two faces containing e must lie in a nontrivial 3-chain, F and an appropriate portion of R_u^{k+1} determine a cellular subcomplex $C \subset A_{k+1}$ with $C \cap R_u^{k+2} = \emptyset$. This gives rise to the same sort of contradictions which are obtained in the proof of Lemma 7. □

COROLLARY 3. *If G is noninductive with respect to vertices u and v, and $\lambda(u, v) = n$, then G consists of $n-1$ concentric cycles of length 6, each joined to the next by 3 edges in such a way as to preserve 3-connectedness, and capped off top and bottom by the stars of u and v.*

PROOF. Apply Lemmas 4–8. □

4. Drawings of noninductive graphs

Let G be a planar cubic 3-connected graph noninductive with respect to u and v, and let D be a drawing of G on the sphere such that none of the hexagonal cycles R_u^k, $1 \leqq k \leqq \lambda(u, v) - 1$ are crossed. Let G' be obtained from G by replacing each of the vertices u and v by a triangular cycle, and let D' be a drawing of G' obtained from D by replacing u and v with triangles sufficiently tiny to insure that none of the R_u^k's nor the two new triangles are crossed. For convenience, we will refer to the triangle obtained from u as R_u^0 and the triangle obtained from v as R_u^n, where $\lambda(u, v) = n$. In D' the interior of R_u^0 is empty, so by means of the usual puncture and stretch operation, we can consider D' to be drawn on the plane and completely contained within triangle R_u^0.

LEMMA 9. *In D', if R_u^{k+1} is inside R_u^k, then R_u^j is inside R_u^k for $k+1 \leqq j \leqq n$. Also, R_u^j is not inside R_u^k for $j<k$.*

PROOF. The proofs that these two conditions hold are essentially the same, and so we only prove the second.

Suppose R_u^j is inside R_u^k for some $j<k$. Let i be the least integer such that R_u^i is inside R_u^k and $i<k$. $i \neq 1$ by assumption, and so R_u^{i-1} is outside of R_u^k, and thus R_u^k must be crossed by the index $(i-1)$ u-struts, contrary to hypothesis. □

We define the *flight diagram* of D to be an array of the integers 0 through n, arranged according to the following two rules: First, if R_u^{k+1} is outside of R_u^k in D', $k+1$ goes directly below k. Second, if R_u^{k+1} is inside R_u^k in D', $k+1$ goes directly to the right of k. Let $\#(D)$ be the number of columns in the flight diagram, and let c_i for $i=1,\dots,\#(D)$ be the number of entries in the i^{th} column.

LEMMA 10. $\lambda_D(u,v) \geqq \#(D)-1$.

PROOF. Let a_i for $i=1,\dots,\#(D)$ be the bottom entry in the i^{th} column. Each R_u^k which separates u from v in D contributes at least 1 to $\lambda_D(u,v)$. Since the $R_u^{a_i}$ for $i=2,\dots,\#(D)-1$ we have $\lambda_D(u,v) > \#(D)-2$, and thus $\lambda_D(u,v) \geqq \#(D)-1$. □

LEMMA 11. $\operatorname{cr}(D) \geqq \sum_{i=1}^{\#(D)} (c_i-1)$.

PROOF. Let P_m be the subgraph of G' induced by R_u^{m-1}, R_u^m, and R_u^{m+1} for $1 \leqq m \leqq n-1$. Let a *relevant crossing* in D' be a crossing where both edges involved lie in a single P_m for some value of m. Now, if $(m, m+1)$ is a pair of integers such that $m+1$ is below m in the flight diagram, then R_u^{m+1} is not inside R_u^m in D'. Either $m-1$ is above m in the flight diagram, or else $m-1$ is to the left of m in the flight diagram. In either case, P_m contains at least one relevant crossing. Thus there is at least one crossing in D' for each pair of integers $(m, m+1)$ where $m+1$ is below m in the flight diagram. In column i there are exactly c_i-1 such pairs, and thus altogether there are $\sum_{i=1}^{\#(D)} (c_i-1)$ such pairs. Therefore $\operatorname{cr}(D) = \operatorname{cr}(D') \geqq \sum_{i=1}^{\#(D)} (c_i-1)$. □

THEOREM 2. *Let G be a planar cubic 3-connected graph noninductive with respect to vertices u and v. Let D be a plan drawing of G such that none of the R_u^k's are crossed for $k=1,\dots,n-1$. Then $\operatorname{cr}(D)+\lambda_D(u,v) \geqq \lambda(u,v)$.*

PROOF. By Lemma 10, $\lambda_D(u,v) \geqq \#(D)-1$. By Lemma 11, $\operatorname{cr}(D) \geqq$

$\geqq \sum_{i=1}^{\#(D)} (c_i - 1)$. Thus

$$\mathrm{cr}(D) + \lambda_D(u,v) \geqq \sum_{i=1}^{\#(D)} (c_i - 1) + \#(D) - 1 =$$

$$= \Big(\sum_{i=1}^{\#(D)} c_i \Big) - \#(D) + \#(D) - 1 = n + 1 - 1 = n = \lambda(u,v). \qquad \square$$

5. The main theorem

THEOREM 3. *Let G be a planar cubic 3-connected graph, and let u, v be vertices of G. Let D be a plane drawing of G. Then* $\mathrm{cr}(D) + \lambda_D(u,v) \geqq \lambda(u,v)$.

PROOF. We proceed by induction on the number of edges in G. The theorem is clearly true for K_4 since the maximum limbic distance is 1. Suppose now it is true for graphs with n edges and let G have $n+1$ edges. Now, if G has a removable edge e which is not limbic with respect to u and v, let $G' = G - e$, and let D' be the drawing of G' obtained from D by removing e. Then by the inductive hypothesis, $\mathrm{cr}(D') + \lambda_{D'}(u,v) \geqq \lambda_{G'}(u,v)$. But $\mathrm{cr}(D) \geqq \mathrm{cr}(D')$, $\lambda_D(u,v) \geqq \lambda_{D'}(u,v)$, and $\lambda_{G'}(u,v) = \lambda_G(u,v)$ since e is not limbic with respect to u and v. Thus $\mathrm{cr}(D) + \lambda_D(u,v) \geqq \lambda_G(u,v)$.

On the other hand, if all removable edges of G are limbic with respect to u and v, then G is noninductive with respect to u and v. If D has none of the R_u^k's crossed, then Theorem 3 holds by dint of Theorem 2. If, however, one of the R_u^k's is crossed, let $e \in R_u^k$ be a crossed edge. Edge e is removable, since all the edges in the R_u^k's are, and we let G' and D' be as above. Again by the inductive hypothesis, $\mathrm{cr}(D') + \lambda_{D'}(u,v) \geqq \lambda_{G'}(u,v)$. Now, since e is crossed in D, $\mathrm{cr}(D) > \mathrm{cr}(D')$. As above, $\lambda_D(u,v) \geqq \lambda_{D'}(u,v)$, and now $\lambda_{G'}(u,v) = \lambda_G(u,v) - 1$. Thus $\mathrm{cr}(D) + \lambda_D(u,v) > \lambda_G(u,v) - 1$, and so $\mathrm{cr}(D) + \lambda_D(u,v) \geqq \lambda_G(u,v)$. $\square$

THEOREM 4. *Let G be a planar cubic 3-connected graph, and let u and v be nonadjacent vertices of G. Then* $\mathrm{cr}(G + uv) = \lambda(u,v) - 1$.

PROOF. The obvious drawing suffices to show that $\mathrm{cr}(G+uv) \leqq \lambda(u,v) - 1$. To reverse the inequality, we invoke Theorem 3. Let E be any drawing of $G + uv$. Let D be a drawing of G obtained from E by removing uv. Then clearly $\mathrm{cr}(E) = \mathrm{cr}(D) + \lambda_D(u,v) - 1$, and applying Theorem 3 to D yields $\mathrm{cr}(D) + \lambda_d(u,v) \geqq \lambda(u,v)$. Therefore $\mathrm{cr}(E) - \lambda_D(u,v) + 1 + \lambda_D(u,v) \geqq \lambda(u,v)$, from which the desired result follows immediately. $\square$

6. Conclusion and acknowledgements

Naturally we conjecture that Theorem 4 holds even if the word "cubic" is deleted from the hypotheses. We conjecture further that if a graph G has a unique embedding on a surface S, then the natural extension of Theorem 4 to drawings of $G+e$ on S holds as well. We would like to acknowledge the influence which the arguments of Ringeisen and Beineke [2, 5] had on our proof of the crucial Lemma 11.

REFERENCES

[1] Barnette, D. W. and Grünbaum, B., On Steinitz's theorem concerning convex 3-polytopes and on some properties of planar graphs, *The Many Facets of Graph Theory* (Proc. Conf., Western Mich. Univ., Kalamazoo, Mich., 1968), Lecture Notes in Math., Vol. 110, Springer, Berlin–New York, 1969, 27–40. *MR* **40** #4148

[2] Beineke, L. W. and Ringeisen, R. D., On the crossing numbers of products of cycles and graphs of order four, *J. Graph Theory* **4** (1980), 145–155. *MR* **81i**:05059

[3] Chartrand, G. and Lesniak, L., *Graphs and digraphs*, 2nd ed., The Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA, 1986. *MR* **87h**:05001

[4] Harary, F., Kainen, P. C. and Schwenk, A. J., Toroidal graphs with arbitrarily high crossing numbers, *Nanta Math.* **6** (1973), 58–67. *MR* **49** #103

[5] Ringeisen, R. D. and Beineke, L., The crossing number of $C_3 \times C_n$, *J. Combinatorial Theory Ser. B* **24** (1978), 134–136. *MR* **80d**:05027

[6] Riskin, A., Polyhedral maps on the pinched torus, Dissertation, University of California, Davis, CA, 1989.

[7] Steinitz, E., Polyeder und Raumeinteilungen, *Encyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen*, Band III, Geometrie, 1. Teil, 2. Hälfte, Heft 9, 12 (III A B 12), B. G. Teubner, Leipzig, 1922, 1–139.

[8] Whitney, H., Congruent graphs and the connectivity of graphs, *Amer. J. Math.* **54** (1932), 150–168. *Jb. Fortschritte Math.* **58**, 608–609. *Zbl* **3**, 328

(Received November 12, 1991)

DEPARTMENT OF MATHEMATICS
NORTHERN ARIZONA UNIVERSITY
P.O. BOX 5717
FLAGSTAFF, AZ 86011
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 415–421

AN INVERSE-FUNCTION THEOREM IN TOPOLOGICAL GROUPS

B. SLEZÁK

The inverse-function theorem of Graves has several generalizations. The main problem is to give a good enough topology on the linear space $\mathcal{L}(X,Y)$ and thus to define the continuity of the derivative. We cannot give a good topology as the norm-topology in Banach-space case. Counterexamples can be found for Fréchet spaces for instance in [1] and [2], when the inverse-function theorem is not valid. However, there are given inverse-function theorems for Fréchet spaces in [3], for sequentially complete locally convex spaces in [4].

However, there is another possibility. Bourbaki has given an inverse-function theorem in Banach spaces under weaker assumptions in [5], namely instead of the continuously differentiability it is enough to suppose that the function is strictly differentiable. Inverse-function theorems using some generalizations of the strictly differentiability for F-spaces can be found in [6] and [8], and for locally convex spaces in [7].

In the first part of this article the Graves theorem ([5]) and the Milutin theorem ([6]) have been compared (the Milutin theorem is much more general and stronger one). In the second part the condition of the strictly differentiability has been generalized in a different way from that in [6], and under this condition some inverse-function theorems have been proved for topological groups. Especially in Banach space case a stronger theorem than the Graves one exists, in F-space case a theorem which is analogous to Milutin's theorem has been obtained.

I

DEFINITION 1. Let X and Y be Banach spaces, $U \subset X$, $f\colon U \to Y$ a function, $u \in U$, $B(u;\delta) := \{x \mid x \in X, \|x-u\| \leqq \delta\}$. The function f is *strictly differentiable* at u if U is a neighbourhood of u and $\forall \varepsilon > 0, \exists \delta > 0, \forall x, y \in \in B(u;\delta) \subset U$:

$$\|(A-f)(x) - (A-f)(y)\| \leqq \varepsilon \|x-y\|,$$

that is if for every positive number ε there is a neighbourhood $B(u;\delta)$ that $A - f$ satisfies the Lipschitz condition on $B(u;\delta)$ with the constant ε. □

1991 *Mathematics Subject Classification.* Primary 22A10; Secondary 46G05.

Key words and phrases. Topological group, inverse function, homeomorphism.

0081–6906/96/$ 5.00 ©1996 *Akadémiai Kiadó, Budapest*

Theorem 1 (Graves theorem, [5]). *Let the function $f\colon U \to Y$ be strictly differentiable at u.*

(i) *If $Df(u)\colon X \to Y$ is an isomorphism, then $f\colon U \to Y$ is locally homeomorphism, that is $\exists B(u;\delta)\colon f|_{B(u;\delta)}$ is a homeomorphism.*

(ii) *If $Df(u)\colon X \to Y$ is surjective, then $\exists B(u;\delta)\colon f|_{B(u;\delta)}$ is an open mapping.*

Definition 2. (i) A family Σ of closed balls in the metric space (X,d) is called a *complete system* if $B(x;r) \in \Sigma$ implies $B(x';r') \in \Sigma$, whenever $r' + d(x,x') \leqq r$.

Let a and b be positive numbers, X and Y metric spaces, and $T\colon X \to Y$ a function.

(ii) T is said to be an *a-covering* on the system Σ if $\forall B(x;r) \in \Sigma :$ $B(T(x); ar) \subset T(B(x;r))$.

(iii) T is *b-compressed* on Σ if $\forall B(x;r) \in \Sigma : T(B(x;r)) \subset B(T(x); br)$. □

It is clear that if T is b-compressed on Σ then T satisfies Lipschitz-condition with the constant b on every ball $B(x;r)$ if $B(x;3r) \in \Sigma$.

Theorem 2 (Milutin, see [6]). *Let (X,d_1) be a metric space, Y a metrizable topological vector space with the translation invariant metric d_2, Σ a complete system in X, $W := \bigcup\limits_{U\in\Sigma} U$, $A\colon X \to Y$ and $f\colon W \to Y$ functions. Suppose that A is continuous on W and a-covering on Σ, $A - f$ is b-compressed on Σ, $0 < b < a$, X is complete. Then f is $(a-b)$-covering on Σ.*

First of all it is shown that if $A\colon X \to Y$ is a surjective continuous linear mapping between Banach spaces then A is a-covering on $\Sigma := \{B(x;r) \mid x \in X, r > 0\}$. As A is open there exists a positive number a such that $B(0;a) \subset A(B(0;1))$. This implies that $\forall r > 0\ \forall x \in X\colon\ B(A(x); ar) \subset A(B(x;r))$. If the function f is strictly differentiable at the point u and $Df(u)$ is surjective then $\forall b > 0$, $\exists B(u;\delta)$ such that $A - f$ satisfies the Lipschitz condition on $B(u;\delta)$ with the constant b, thus $A - f$ is b-compressed on $\Sigma := \{B(x;r) \mid B(x;r) \subseteqq B(u;\delta)\}$. The number δ can be found so that $b < a$. Using the Milutin theorem we get that the function f is *uniformly* open on $B(u;\delta)$. Thus under weaker assumptions (it is supposed only that $b < a$ and that $Df(a)$ is a covering, but its linearity is not used) a stronger statement is obtained than Graves's one.

II

If the conditions of the Milutin theorem are satisfied then $\forall B(x;r) \in \Sigma$:

$$(A-f)(B(x;r)) \subset B((A-f)(x); br) =$$
$$= B(A(x); br) - f(x) \subset A\left(B\left(x; \frac{b}{a} r\right)\right) - f(x) = A(B(x;kr)) - f(x),$$

where $k := \frac{b}{a} < 1$. The assumption $(A - f)(B(x;r)) \subset A(B(x;kr)) - f(x)$ is weaker than the conditions of the Milutin theorem. It is possible that a continuous linear mapping between topological linear spaces is open but it is not a covering. A very simple example is given here: $\mathrm{id}_{\mathbb{R}} : (\mathbb{R}, d_1) \to \to (\mathbb{R}, d_2)$, where $d_1(x,y) := |x-y|$ and $d_2(x,y) := \sqrt{|x-y|}$. The function $\mathrm{id}_{\mathbb{R}}$ is obviously open, linear but not a covering.

In the following theorems only the additivity of A is supposed but the function A is not required to be a covering.

Lemma 1. *Let X be a topological space with a continuous quasimetric d and let Y be a group, $A\colon X \to Y$ and $f\colon X \to Y$ are functions, $x_0 \in \in X$. If A is injective and there exists a number $k \in]0,1[$ such that for $\forall r > > 0$: $(A-f)(B(x_0;r)) \subset A(B(x_0;kr)) - f(x_0)$, then $\forall x \in X$: $d(x,x_0) \neq 0 \Rightarrow \Rightarrow f(x) \neq f(x_0)$.*

Proof. Suppose that $f(x) = f(x_0)$. Then

$$x \in B(x_0; d(x_0,x)) \Rightarrow (A-f)(x) \in A(B(x_0; kd(x_0,x))) - f(x_0) \Rightarrow$$
$$\Rightarrow \exists z \in B(x_0; kd(x_0,x)) \colon (A-f)(x) = A(z) - f(x_0) \Rightarrow A(z) = A(x).$$

But $k < 1 \Rightarrow x \notin B(x_0, kd(x_0,x)) \Rightarrow x \neq z \Rightarrow A(x) \neq A(z)$, which is a contradiction. □

Lemma 2. *Let X be a topological space, d a quasimetric which defines the topology of X. Let Y be a topological group, $A\colon X \to Y$ and $f\colon X \to Y$ be functions. If there exists a positive number k so that for every positive number r the inclusion $(A-f)(B(x_0;r)) \subset A(B(x_0;kr)) - f(x_0)$ holds and A is continuous at x_0 then f is also continuous at x_0.*

Proof. Denote $B_Y(0)$ a basis of the neighbourhoods of zero in Y with symmetric elements. Let $W \in B_Y(0)$ be arbitrary, $W' \in B_Y(0)$, $W' + W' \subset \subset W$. As A is continuous at x_0 the number $r > 0$ can be chosen so that $A(B(x_0;r)) \subset A(x_0) + W'$ and $A(B(x_0;kr)) \subseteqq A(x_0) + W$. If $u \in B(x_0;r)$ then

$$(A-f)(u) \in (A-f)(B(x_0;r)) \subset A(B(x_0;kr)) - f(x_0),$$

hence $\exists z \in B(x_0;kr)$: $A(u) - f(u) = A(z) - f(x_0)$

$$f(u) = f(x_0) - A(z) + A(u) \in f(x_0) - A(B(x_0;kr)) + A(B(x_0;r)) \subset$$
$$\subset f(x_0) - W' - A(x_0) + A(x_0) + W' = f(x_0) + W' + W' \subset f(x_0) + W.$$

This means that $f(B(x_0;r)) \subset f(x_0) + W$. □

Lemma 3. *Let G be a topological group, d a continuous translation invariant quasimetric on G, $X \subset G$ a set, Y a group, $k \in]0,1[$. If the function $A\colon G \to Y$ is additive, $f\colon X \to Y$ is a function such that for*

$$\forall B(x;r) \subset X\colon \ (A-f)(B(x;r)) \subset A(B(x;kr)) - f(x)$$

then $\forall B(X;r) \subset X$, $\forall y \in f(x) + A(B(0;(1-k)r))$, $\exists \{x_n\} \subset X : \{x_n\}$ *is a Cauchy sequence (with respect to* d*), furthermore for* $\forall n \in N$: $A(x_{n+1}) = (A-f)(x_n) + y$.

PROOF. Let $x_0 := x$. Then

$$(A-f)(x_0)+y \in (A-f)(x_0)+f(x_0)+A(B(0;(1-k)r)) = A(B(x_0;(1-k)r)).$$

Hence $\exists x_1 \in B(x_0;(1-k)r)$: $(A-f)(x_0)+y = A(x_1)$ and $d(x_1,x_0) \leqq (1-k)r \Rightarrow x_1 \in B(x_0;r)$. Now it will be shown by induction that

$$\forall n \in N, \exists x_{n+1} \in B(x_n; k^n(1-k)r) : (A-f)(x_n)+y = A(x_{n+1}).$$

Suppose that the statement holds for $m \in N$ if $1 \leqq m \leqq n$. I.e.,

$$\begin{gathered} y = f(x_{n-1}) - A(x_{n-1}) + A(x_n), \\ (A-f)(x_n) + y \in (A-f)(B(x_{n-1}; k^{n-1}(1-k)r)) + \\ + f(x_{n-1}) + A(-x_{n-1}+x_n) \subset \\ \subset A(B(x_{n-1}; k^n(1-k)r)) + A(-x_{n+1}+x_n) = A(B(x_n; k^n(1-k)r)). \end{gathered}$$

Hence

$$\exists x_{n+1} \in B(x_n; k^n(1-k)r) : (A-f)(x_n) + y = A(x_{n+1}).$$

If $0 \leqq m \leqq n$ then

$$d(x_{n+1}, x_m) \leqq \sum_{i=m}^{n} k^i(1-k)r = k^m r(1-k) \sum_{j=0}^{n-m} k^j \leqq k^m r \leqq r,$$

consequently, $x_{n+1} \in B(x;r)$ and $\{x_n\}$ is a Cauchy sequence with respect to d. □

THEOREM 3. *Let* $\mathcal{G}$ *be a topological group with the translation invariant metric* d, Y *a topological group,* $X \subset \mathcal{G}$ *an open set,* $f : X \to Y$ *and* $A : \mathcal{G} \to Y$ *functions,* A *be continuous,* $k \in]0,1[$. *If* $\forall B(x;r) \subseteqq X$: $(A-f)(B(x;r)) \subseteqq A(B(x;kr)) - f(x)$ *then*

(i) *The function* f *is continuous. This continuity is uniform on every closed ball in* X.

(ii) *If* A *is injective then* f *is injective on every closed ball* $B(x_0;r)$ *if* $B(x_0;3r) \subset X$.

Furthermore, if the topology of $\mathcal{G}$ *is complete and* Y *has a Hausdorff topology,* A *is additive, then*

(iii) $\forall B(x;r) \subset X$: $f(x) + A(B(0;(1-k)r)) \subset f(B(x;r))$.

(iv) *If* A *is open then* f *is uniformly open on the set* $X_r := \{x \in X \mid B(x;r) \subset X\}$ *for every* $r > 0$.

(v) $\forall x \in X$, $\exists f_r : \mathrm{im}(f) \to X$, *the function* f_r *is a right-inverse of* f *and for*

$$\forall r > 0 : B(x;r) \subset X \Rightarrow f_r(f(x) + A(B(0;(1-k)r))) \subset B(x;2r).$$

It means that f_r is continuous at $f(x)$ and $f_r(f(x)) = x$.

(vi) *If A is an isomorphism then f is a homeomorphism on the balls $B(x;r)$, moreover $f|_{B(x;r)}$ and $f^{-1}|_{f(B(x;r))}$ are uniformly continuous, whenever $B(x;3r) \subset X$.*

PROOF. (i) It follows from Lemma 2 that if W is a neighbourhood of the zero element of Y then there exists a positive number r such that $f(B(x;r)) \subset \subset f(x) + W$, whenever $B(x;r) \subset X$ and $A(B(x;r)) \subset A(x) + W'$, $W' + W \subset \subset W$. If $A(B(0;r)) \subset W'$ then $B(x;r) \subset X \Rightarrow A(B(x;r)) = A(x + B(0;r)) \subset \subset A(x) + W'$, hence r does not depend on x.

(ii) Let $B(x_0;3r) \subset X$. Then for $\forall x \in B(x_0;r)$: $B(x;r) \subset B(x_0;2r)$ and $B(x_0;r) \subset B(x;2r) \subset B(x_0;3r)$.

By Lemma 1 if $u \in B(x_0;r)$ and $u \neq x$ then $f(x) \neq f(u)$.

(iii) Using Lemma 3 we have a Cauchy sequence $\{x_n\} \subset B(x;r)$ such that for $\forall n \in \mathbb{N}$: $(A - f)(x_n) + y = A(x_{n+1})$, if $y \in f(x) + A(B(0;(1-k)r))$. As the topology of $\mathcal{G}$ is complete there exists a limit x' of $\{x_n\}$. A and f are continuous, therefore

$$(A - f)(x_n) + y \to A(x') - f(x') + y = A(x')$$

as the topology of Y is of Hausdorff type. Hence $y = f(x')$ and $f(x) + + A(B(0;(1-k)r)) \subset f(B(x;r))$.

(iv) If A is open then the set $A(B(0;(1-k)r))$ is also open and $f(B(x;r))$ is a neighbourhood of $f(x)$. If $x \in X_r$ then the neighbourhood $A(B(0;(1 - - k)r))$ does not depend on x.

(v) Let $y \in f(x) + A(B(0;(1-k)r))$. Denote $S := \inf\{s \mid -f(x) + y \in \in A(B(0;(1-k)s))\}$. Obviously, $S \leqq r$ and $y \in f(x) + A(B(0;(1-k)2r)) \subseteqq f(B(x;2r))$, thus $\exists u \in B(x;2r)$: $f(u) = y$. Let be $f_r(y) := u$. Then $f_r(f(x) + + B(0;(1-k)r)) \subset B(x;2r)$ and

$$f_r(f(x)) \in \bigcap_{r>0} B(x;2r) = \{x\}.$$

(vi) It is clear that f is a homeomorphism on $B(x;r)$. It must be shown only that f^{-1} is uniformly continuous. It follows from (v) that $f^{-1}(f(x) + + A(B(0;(1-k)r))) \subset B(x;2r)$, where the neighbourhood $A(B(0;(1-k)r))$ does not depend on $f(x)$. □

THEOREM 4. *Let $\mathcal{G}$ be a not necessarily metrizable topological group, $\{d_i \mid i \in I\}$ the family of translation invariant quasimetrics determining the Hausdorff type topology of $\mathcal{G}$, X a subset of $\mathcal{G}$ and $x_0 \in X$, Y be a group, $A\colon \mathcal{G} \to Y$ a bijective and additive mapping, $f\colon X \to Y$ a function. If for $\forall i \in I$, $\exists k_i \in]0;1[$, $\forall x \in X$, $\forall B_i(x;r)$:*

$$(A - f)(B_i(x;r) \cap X) \subset A(B_i(x;k_i r) \cap X) - f(x)$$

then

(i) *f is injective;*

(ii) *$f:(X,\tau)\to(Y,\tau')$ is a continuous mapping, where τ is the relative topology of X and τ' is the topology of Y defined by A;*

(iii) *If the topology of $\mathcal{G}$ is sequentially complete and of Hausdorff type, $J\subseteqq I$, then for*

$$\forall r>0,\ \forall i_0\in I,\ \forall x\in X:\ \bigcap_{j\in J} B_j(x;r_j)\subset X\Rightarrow$$
$$\Rightarrow f(x)+A(\bigcap_{j\in J} B_j(0;r_j(1-k_j))\cap B_{i_0}(0;r(1-k_{i_0})))\subset f(X\cap B_{i_0}(x;r)).$$

Moreover, if Y is a topological group and A is open, then f is open too, therefore it is homeomorphism.

PROOF. First of all under these conditions a topology can be defined on Y by A so that A is an isomorphism.

(i) As $x,u\in X$ and $x\neq u$, there exists a quasimetric d_i such that $d_i(x,u)\neq 0$. By Lemma 1, $f(x)\neq f(u)$.

(ii) Let $x\in X$, W an arbitrary neighbourhood of the zero element of Y ($W\in B_Y(0)$), $W'\in B_Y(0)$ a symmetric neighbourhood, $W'+W'\subset W$. As A is continuous there exists a finite subset J of I such that $A(\bigcap_{i\in J} B_i(x;r))\subset$ $\subset A(x)+W'$.

$$A(\bigcap_{i\in J} B_i(x;r)\cap X)\subset A(x)+W'$$

and

$$A(\bigcap_{i\in J} B_i(x;k_ir)\cap X)\subset A(x)+W'.$$

If $u\in\bigcap_{i\in J} B_i(x;r)\cap X$ then

$$(A-f)(u)\in(A-f)(\bigcap_{i\in J} B_i(x;r)\cap X)\subset\bigcap_{i\in J}(A(B_i(x;k_ir)\cap X)-f(x))=$$
$$=\bigcap_{i\in J} A(B_i(x;k_ir)\cap X)-f(x).$$

$$(A-f)(u)+f(x)\in A(\bigcap_{i\in J} B_i(x;k_ir)\cap X),$$

hence

$$f(u)\in f(x)-A(\bigcap_{i\in J} B_i(x;k_ir)\cap X)+A(\bigcap_{i\in J} B_i(x;r)\cap X).$$

$$f(u)\in f(x)-W'-A(x)+A(x)+W'=f(x)+W'+W'\subset f(x)+W.$$

Hence $f(\bigcap_{i\in J} B_i(x;r)\cap X)\subset f(x)+W$.

(iii) Let $x \in \bigcap_{j \in J} B_j(x; r_j) \subset X$. Denote $H := J \cup \{i_0\}$. Suppose that $y \in \in f(x) + A(\bigcap_{h \in H} B_h(0; (1-k_h)r_h))$. Using Lemma 3 for the topological group $(\mathcal{G}, d_h)$, a sequence $\{x_n\} \subset \bigcap_{h \in H} B_h(x; r_h)$ can be chosen which is Cauchy sequence with respect to every d_h, $h \in H$, and $\forall n \in \mathbb{N}$: $(A-f)(x_n) + y = A(x_{n+1})$. (As A is injective we have the same $\{x_n\}$ for every $h \in H$.) Now let $i \in I$ be arbitrary. $\exists r' > 0$; $y \in f(x) + A(X \cap B_i(x; (1-k)r'))$. Using Lemma 3 again it is obtained that $\{x_n\}$ is a Cauchy sequence in $\mathcal{G}$. $\mathcal{G}$ is sequentially complete, thus $\{x_n\}$ is convergent. Denote $x' := \lim x_n \in \bigcap_{h \in H} B(x; r_h)$. Then $(A-f)(x') + y = A(x')$ and $f(x') = y$, because the topology of Y is of Hausdorff type. Therefore $f(x) + A(\bigcap_{h \in H} B_h(0; (1-k_h)r_h)) \subseteq f(\bigcap_{h \in H} B_h(x; r_h))$. It follows that if A is an open mapping then f is also open on X since in this case the topology of Y is finer than the topology defined by A. Consequently, f is homeomorphism. □

REFERENCES

[1] Lojasiewicz, S. Jr., An example of continuous injective polynomial with nowhere dense range, whose differential at each point is an isomorphism, *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **24** (1976), 1109–1111. *MR* **55** #11292

[2] Averbuh, V. I. and Smoljanov, O. G., Different definitions of derivative in linear topological spaces, *Uspehi Mat. Nauk* **23** (1968), No. 4 (142), 67–116 (in Russian). *RŽMat* **1969** #4 B480; *MR* **39** #7424

[3] Lojasiewicz, S. Jr. and Zehnder, E., An inverse function theorem in Fréchet-spaces, *J. Funct. Anal.* **33** (1979), 165–174. *MR* **80k**:58014

[4] Suhinin, M. F., The local invertability of a differentiable mapping, *Uspehi Mat. Nauk* **25** (1970), no. 5, 249–250 (in Russian). *MR* **44** #5776b

[5] Bourbaki, N., *Éléments de mathématique*. Fasc. XXXIII, *Variétés différentielles et analytiques*. Fascicule de résultats (Paragraphes 1 à 7), Actualités scientifiques et industrielles, no. 1333, Hermann, Paris, 1967. *MR* **36** #2161

[6] Dmitruk, A. V., Milyutin, A. A. and Osmolovskiĭ, N. P., Ljusternik's theorem and the theory of the extremum, *Uspekhi Mat. Nauk* **35** (1980), 11–46 (in Russian). *MR* **82c**:58010

[7] Lorenz, K. J., An implicit function theorem in locally convex spaces, *Math. Nachr.* **129** (1986), 91–101. *MR* **88c**:58008

[8] Slezák, B., On the inverse function theorem in F-spaces, *Studia Sci. Math. Hungar.* (to appear).

(Received September 3, 1991)

VESZPRÉMI EGYETEM
MATEMATIKAI ÉS SZÁMÍTÁSTUDOMÁNYI TANSZÉK
POSTAFIÓK 158
H-8201 VESZPRÉM
HUNGARY

slezakb@almos.vein.hu

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 423–430

NOTES ON QUANTUM ENTROPY

M. OHYA and D. PETZ

The present paper consists of two parts. In the first one it will be proved that the von Neumann entropy governs the size of rather sure projections in the course of independent trials. The second part is devoted to the extension of the von Neumann entropy to states of arbitrary unital C^*-algebras.

By a finite quantum system we mean an algebra of matrices which is stable under taking adjoint. (In other words, a finite quantum system is a finite dimensional C^*-algebra.) If $\mathcal{A}$ is such an algebra then there is a linear functional Tr which takes the value 1 at each minimal projection. It is "tracial" in the sense that

$$\operatorname{Tr} ab = \operatorname{Tr} ba \qquad (a, b \in \mathcal{A}).$$

Every functional ω on $\mathcal{A}$ is determined by a density operator $D_\omega \in \mathcal{A}$ in the form

$$\omega(a) = \operatorname{Tr} D_\omega a \qquad (a \in \mathcal{A}).$$

The entropy $S(\omega)$ of a functional ω is defined by means of its density operator as

$$S(\omega) = \operatorname{Tr} \eta(D_\omega).$$

This notion was introduced by von Neumann in 1927 and we term it von Neumann's entropy or shortly entropy (cf. [8]).

It is understood in probability theory that the notion of (Shannon or measure theoretic) entropy has successful applications in a variety of subjects because it determines the asymptotic behaviour of certain probabilities in the course of independent trials (see, for example [1] or [11]). Now we will discuss this phenomenon for finite quantum systems.

Let $\mathcal{A}$ be a finite quantum system with a faithful state ω. The n-fold algebraic tensor product $\mathcal{A}_n = \mathcal{A} \otimes \ldots \otimes \mathcal{A}$ is again a finite quantum system and the product functional $\omega_n = \omega \otimes \ldots \otimes \omega$ is a state of $\mathcal{A}_n$. Using the obvious identifications the inclusion $(\mathcal{A}_n, \omega_n) \subset (\mathcal{A}_m, \omega_m)$ holds for $n \leqq m$ and we set

$$(\mathcal{A}_\infty, \omega_\infty) = \cup\{(\mathcal{A}_n, \omega_n) : n \in \mathbb{N}\}.$$

1991 *Mathematics Subject Classification.* Primary 46L30; Secondary 82B10, 94A17.

Key words and phrases. Density matrix, state space, von Neumann entropy, relative entropy.

On the $*$-algebra $\mathcal{A}_\infty$ the right shift endomorphism γ is defined for $a_1 \otimes a_2 \otimes \otimes \ldots \otimes a_n \in \mathcal{A}_n$ as

$$\gamma(a_1 \otimes a_2 \otimes \ldots \otimes a_n) = I \otimes a_1 \otimes a_2 \otimes \ldots \otimes a_n \in \mathcal{A}_{n+1}$$

and ω_∞ is invariant under γ. Now perform the GNS-constructions with the state ω_∞ and arrive at the triplet $(\pi, \mathcal{H}, \Omega)$. We identify $\mathcal{A}_\infty$ through its faithful representation π with a subalgebra of the generated von Neumann algebra $\mathcal{M} = \pi(\mathcal{A}_\infty)'' \subset \mathcal{B}(\mathcal{H})$. The normal state

$$\omega(a) = \langle \Omega, a\Omega \rangle \qquad (a \in \mathcal{M})$$

is an extension of ω_∞ and the endomorphism γ extends to $\mathcal{M}$ such that the relation $\omega \circ \gamma = \omega$ is preserved. (For the sake of simpler notation we do not use a new letter for the extension.)

The following result may be called the weak law of large numbers (for independent finite quantum systems). Since it is well known its proof will be omitted.

PROPOSITION 1. *In the above described situation the following statements hold.*

(i) *If $a \in \mathcal{M}$ and $\gamma(a) = a$ then $a \in \mathbf{C}I$.*

(ii) *For every $a \in \mathcal{M}$ the sequence $s_n(a) = n^{-1}(a + \gamma(a) + \ldots + \gamma^{n-1}(a))$ converges to $\omega(a)I$ in the strong operator topology.*

(iii) *If $a \in \mathcal{M}^{sa}$ and $J \subset \mathbf{R}$ is closed interval such that $\omega(a) \notin J$ and p_n is the spectral projection of $s_n(a)$ corresponding to the interval J then $p_n \to 0$ in the strong operator topology.*

Let us fix a positive number $\varepsilon < 1$. For a while we say that a projection $Q_n \in \mathcal{A}_n$ is rather sure if $\omega_n(Q_n) \geqq 1 - \varepsilon$. On the other hand, the size of Q_n, the cardinality of a maximal pairwise orthogonal family of projections contained in Q_n, is given by $\mathrm{Tr}_n Q_n$. (The subscript n in Tr_n indicates that the algebraic trace functional on $\mathcal{A}_n$ is meant here.) The theorem below says that the von Neumann's entropy of ω governs asymptotically the size of rather sure projections: A rather sure projection in $\mathcal{A}_n$ contains at least $\exp(nS(\omega))$ pairwise orthogonal minimal projection.

THEOREM 2. *Under the above conditions and with the above notation the limit relation*

$$\lim_{n\to\infty} \frac{1}{n} \inf \{\log \mathrm{Tr}_n Q_n : Q_n \in \mathcal{A}_n \text{ is a projection, } \omega_n(Q_n) \geqq 1 - \varepsilon\} = S(\omega)$$

holds.

PROOF. If D_n denotes the density of ω_n then one can see easily that

$$-\log D_n = \sum_{i=0}^{n-1} \gamma^i(-\log D_1)$$

where γ stands for the right shift. The sequence $(\gamma^i(-\log D_1))$ behaves as independent identically distributed random variables with respect to the state ω_∞. More precisely, the previous proposition applies for $a = -\log D_1$ and tells

$$\frac{1}{n}\log D_n \to S(\omega)I$$

strongly. Let $P(n,\delta)$ be the spectral projection of the selfadjoint operator $-n^{-1}\log D_n$ corresponding to the interval $(S(\omega)-\delta, S(\omega)+\delta)$. According to (iii) of Proposition 1 one has

$$P(n,\delta) \to I \tag{1}$$

strongly for every $\delta > 0$. In particular,

$$\omega(P(n,\delta)) = \langle P(n,\delta)\Omega, \Omega\rangle \to 1$$

as $n \to \infty$ and $P(n,\delta)$ is a rather sure projection if n is large enough. It follows from the definition of $P(n,\delta)$ that

$$D_n P(n,\delta)\exp(nS(\omega)-n\delta) \leqq P(n,\delta) \leqq D_n \exp(nS(\omega)+n\delta) \tag{2}$$

which gives

$$\frac{1}{n}\log \mathrm{Tr}_n P(n,\delta) \leqq S(\omega)+\delta.$$

Since $\delta > 0$ was arbitrary we establish

$$\limsup_{n\to\infty} \frac{1}{n}\inf\{\log \mathrm{Tr}_n Q_n : Q_n\} \leqq S(\omega). \tag{3}$$

To prove that $S(\omega)$ is actually the limit we shall argue by contradiction. Assume that there exist a sequence $n(1) < n(2) < \dots$ of integers, a number $t > 0$ and projections $Q(n(k)) \in \mathcal{A}_{n(k)}$ $(k = 1,2,\dots)$ such that

(i) $\omega_\infty(Q(n(k)) \geqq 1-\varepsilon$,

(ii) $\log \mathrm{Tr}_{n(k)} Q(n(k)) \leqq n(k)(S(\omega)-t)$.

The bounded sequence $(Q(n(k)))_k$ has a weak limit point in the von Neumann algebra $\mathcal{M}$, say $T \in \mathcal{M}$. Instead of selecting a subsequence we suppose that $Q(n(k)) \to T$ weakly. It is straightforward to show that from (1) the weak limit

$$Q(n(k))P(n(k),\delta) \to T$$

follows. Consequently,

$$\liminf_{k\to\infty} \omega_\infty(Q(n(k))P(n(k),\delta)) \geqq \omega(T) \geqq 1-\varepsilon. \tag{4}$$

Using the first part of (2) we estimate

$$\begin{aligned}\mathrm{Tr}\, Q(n(k)) &\geqq \mathrm{Tr}\, Q(n(k))P(n(k),\delta)\\ &\geqq \mathrm{Tr}\, D_{n(k)}Q(n(k))P(n(k),\delta)\exp(nS(\omega)-n\delta)\\ &= \exp(nS(\omega)-n\delta)\omega_\infty(Q(n(k))P(n(k),\delta))\end{aligned}$$

and

$$\liminf_{k\to\infty} \frac{1}{n(k)} \log \mathrm{Tr}_{n(k)} Q(n(k))$$
$$\geqq S(\omega) - \delta + \lim_{k\to\infty} \frac{1}{n(k)} \log \omega_\infty(Q(n(k))P(n(k),\delta)).$$

The limit term on the right-hand side vanishes due to (4) and we arrive at a contradiction with (ii) if $0<\delta<t$. This proves the theorem.

Opposite to the commutative case the state space of a quantum system is not a Choquet simplex in the sense that states admit several extremal decompositions. For example, for $\mathcal{A} = M_2(\mathbf{C})$ the general form of a density matrix is

$$D = \frac{1}{2}\begin{pmatrix} 1+a & b+ic \\ b-ic & 1-a \end{pmatrix} \tag{5}$$

where a,b,c are real numbers and $a^2+b^2+c^2 \leqq 1$. Thanks to the affine correspondence $D \leftrightarrow (a,b,c)$ we can visualize the state space as a ball (of radius 1) and surface points correspond to pure states.

Let φ be a state of a finite quantum system and $\varphi = \sum_i \lambda_i \psi_i$ be an extremal decomposition (that is, every ψ_i is pure). Approaching from information theory one might think that the entropy of φ is $-\sum_i \lambda_i \log \lambda_i$. This, however, would not be satisfactory because the λ_i's are not in general the probabilities of mutually exclusive events. In fact,

$$S(\varphi) \leqq -\sum_i \lambda_i \log \lambda_i \tag{6}$$

and the equality holds if and only if the extremal decomposition $\sum \lambda_i \psi_i$ is orthogonal. This was obtained in [4] a long time ago and here it will be deduced by means of the relative entropy. The inequality (6) is interpreted as follows. In the sense of information content, the most economical extremal decomposition is the orthogonal one, which is implemented by the density matrix.

The entropy of ω with respect to φ is defined by

$$S(\omega,\varphi) = \begin{cases} \mathrm{Tr}\, D_\omega(\log D_\omega - \log D_\varphi) & \text{if } \operatorname{supp} D_\varphi \geqq \operatorname{supp} D_\omega \\ +\infty & \text{otherwise.} \end{cases} \tag{7}$$

Here $\operatorname{supp} D_\psi$ denotes the smallest projection p such that $\psi(p) = \psi(I)$. In particular, $S(\omega,\varphi)$ is always finite if the density of φ has strictly positive eigenvalues. (Such a φ is called faithful.) When D_φ commutes with D_ω and their eigenvalue lists are $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ and $(\kappa_1, \ldots, \kappa_n)$, respectively, then $S(.,.)$ reduces to the classical expression due to Kullback and Leibler.

Although we mostly speak of the relative entropy of states it is convenient to allow ω and φ in the definition of $S(\omega,\varphi)$ to be arbitrary positive functionals.

The relative entropy may be defined for linear functionals of an arbitrary C^*-algebra. Now we do not give the details of the rather technical chain of definitions going through von Neumann algebras and normal functionals to an arbitrary C^*-algebra. We just mention a possible extension of (7), the so-called Kosaki's formula ([1], [5]).

$$S(\omega,\varphi)=\sup\sup\left\{\omega(I)\log n-\int_{1/n}^{\infty}[\omega(y(t)^*y(t))+t^{-1}\varphi(x(t)x(t)^*]t^{-1}dt\right\}$$

where the first sup is taken over all natural numbers n, the second one is over all step functions $x\colon(1/n,\infty)\to N$ with finite range and $y(t)=I-x(t)$.

The relative entropy of positive functionals of a C^*-algebra shares the following properties (see [1], [5] and [9]).

(i) $(\omega,\varphi)\mapsto S(\omega,\varphi)$ is convex and weakly lower semicontinuous.

(ii) $\|\varphi-\omega\|^2\leqq 2S(\omega,\varphi)$ if $\varphi(I)=\omega(I)=1$.

(iii) $S(\omega,\varphi_1)\geqq S(\omega,\varphi_2)$ if $\varphi_1\leqq\varphi_2$.

(iv) For a unital Schwarz map $\alpha\colon A_0\mapsto A_1$ the relation $S(\omega\circ\alpha,\varphi\circ\alpha)\leqq$ $\leqq S(\omega,\varphi)$ holds.

(v) For $\omega=\sum_i\omega_i$ we have $S(\omega,\varphi)+\sum_{i=1}^n S(\omega_i,\omega)=\sum_{i=1}^n S(\omega_i,\varphi)$.

Below the properties (ii) and (v) will not be required.

PROPOSITION 3. *Let φ be a state of a finite quantum system $\mathcal{A}$. If φ is a convex combination $\sum_j\mu_j\varphi_j$ of states then*

$$S(\varphi)\geqq\sum_j\mu_j S(\varphi,\varphi_j).$$

When all the φ_j's are pure then the equality sign applies.

PROOF. It suffices to prove the equality because the inequality follows by convexity of the relative entropy. By simple computation we have

$$\sum_j\mu_j S(\varphi,\varphi_j)=\sum_j\mu_j\operatorname{Tr}\eta(D_{\varphi_j})+\operatorname{Tr}\eta(D_\varphi)$$

and the first term on the right-hand side vanishes when all the D_{φ_j}'s are projections.

Now let $\varphi=\sum\lambda_i\psi_i$ be an extremal decomposition. Combining Proposition 3 with the monotonicity (iii) of the relative entropy in the first variable we infer (6) as follows.

$$S(\varphi)=\sum_i\lambda_i S(\varphi,\psi_i)\leqq\sum_i\lambda_i S(\lambda_i\psi_i,\psi_i)=-\sum_i\lambda_i\log\lambda_i.$$

We note that Proposition 3 and its consequence (6) remain valid if $\mathcal{A}$ is a von Neumann algebra which is the direct sum of type I factors and φ is an arbitrary normal state. (In this case the functional Tr in the proof should be understood as the faithful normal semifinite trace which takes the value 1 at each minimal projection.)

Proposition 3 allowed the following definition of the entropy of states of arbitrary C^*-algebras in [7].

$$S(\varphi) = \sup\left\{ \sum_i \lambda_i S(\varphi, \varphi_i) : \sum_i \lambda_i \varphi_i = \varphi \right\}. \tag{8}$$

Here the supremum is taken over all decompositions of φ into finite (or equivalently countable) convex combinations of other states. Apparently the background uniform distribution provided by the trace functional in finite quantum systems is not present in this definition. Some properties of $S(\varphi)$ are immediate from those of the relative entropy. The quantity $S(\varphi)$ is nonnegative and vanishes when and only when, φ is a pure state. Moreover, the entropy is lower semicontinuous because it is the supremum of lower semicontinuous relative entropy functionals (see (i) above). The invariance of $S(\varphi)$ under automorphisms is obvious as well.

PROPOSITION 4. *Let $\mathcal{A}_0 \subset \mathcal{A}$ be C^*-algebras and assume that there exists a conditional expectation of $\mathcal{A}$ onto $\mathcal{A}_0$ preserving a given state φ of $\mathcal{A}$. Then*

$$S(\varphi) \geqq S(\varphi \mid \mathcal{A}_0). \tag{9}$$

PROOF. Indeed, if $\sum \lambda_i \psi_i = \varphi \mid \mathcal{A}_0$ for some states ψ_i of $\mathcal{A}_0$ then $\sum \lambda_i \psi_i \circ E$ is a decomposition of φ where E was written for the conditional expectation in the statement. The rest follows from $S(\varphi \mid \mathcal{A}_0, \psi_i) \leqq S(\varphi, \psi_i \circ E)$.

The existence of the conditional expectation preserving the given state is an essential hypothesis in Proposition 4. The monotonicity property (9) does not hold in general. The simplest counterexample is due to the fact that in the quantum case a pure state of the algebra can yield a restriction which is a mixed state on the subalgebra. The next observation is obvious.

PROPOSITION 5. *Let $\mathcal{A} = \mathcal{A}_1 \oplus \mathcal{A}_2$ be C^*-algebras and $\varphi = \lambda\varphi_1 \oplus (1-\lambda)\varphi_2$ a state of $\mathcal{A}$ $(0 < \lambda < 1)$. Then*

$$S(\varphi) = \lambda S(\varphi_1) + (1-\lambda) S(\varphi_2).$$

Now let φ be a normal state on a von Neumann algebra $\mathcal{M}$. Then a decomposition $\varphi = \sum \lambda_i \varphi_i$ is necessarily built from normal states φ_i if $\lambda_i \neq 0$. Hence, if we wish, in the definition (8) we may restrict ourselves to normal states φ_i. When p is the support projection of φ then $S(\psi, \varphi) = S(\psi \mid p\mathcal{M}p, \varphi \mid p\mathcal{M}p)$ whenever $\psi(p) = 1$ for the state ψ. Consequently

$$S(\varphi) = S(\varphi \mid p\mathcal{M}p). \tag{10}$$

The following result is due ot Hiai ([3]) and its proof uses the structure theory of von Neumann factors.

THEOREM 6. *Let φ be normal state of a von Neumann algebra $\mathcal{M}$ and let p be the support projection of φ. If $p\mathcal{M}p$ is a countable direct sum of type I factors then $S(\varphi) = \operatorname{Tr} \eta(D_\varphi)$ where D_φ is the density of φ with respect to the canonical semifinite normal trace* Tr *on $p\mathcal{M}p$. Otherwise, $S(\varphi) = \infty$.*

Now we turn to entropy of states of C^*-algebras. Let ψ be a state of a C^*-algebra $\mathcal{A}$. We write $\overline{\psi}$ for the vector state induced by the cyclic vector Ψ on the von Neumann algebra $\pi_\psi(\mathcal{A})''$ when (H_ψ, Ψ, π_ψ) is the GNS-triple corresponding to ψ.

LEMMA 7. *With the above notation we have $S(\psi) = S(\overline{\psi})$.*

PROOF. The key to the proof is the fact that a finite relative entropy may be computed in the GNS representation of the reference state. This yields readily that $S(\psi) \geqq S(\overline{\psi})$. On the other hand, if $\psi = \sum_i \lambda_i \psi_i$ is a convex decomposition in the state space of $\mathcal{A}$ then one can find (normal) states $\overline{\psi}_i$ of $\pi_\psi(\mathcal{A})''$ such that $\psi_i = \overline{\psi}_i \circ \pi_\psi$ and $\overline{\psi} = \sum_i \lambda_i \overline{\psi}_i$. Since $S(\psi, \psi_i) = S(\overline{\psi}, \overline{\psi}_i)$ the converse inequality $S(\psi) \leqq S(\overline{\psi})$ follows.

THEOREM 8. *Let ψ be a state of a C^*-algebra $\mathcal{A}$. Then*

$$S(\psi) = \inf \Big\{ \sum_i \eta(\lambda_i) \Big\} \tag{11}$$

where the infimum is taken over all possible decompositions $\psi = \sum_i \lambda_i \psi_i$ into pure states. If ψ is not a countable convex combination of pure states then $S(\psi) = \infty$.

PROOF. Lemma 7 allows us to reduce the theorem to the von Neumann algebra version. We have to consider two cases. If ψ is not a countable convex combination of pure states, then the von Neumann algebra $p\pi_\psi(\mathcal{A})''p$ is not a countable direct sum of type I factors. (p is the support projection of $\overline{\psi}$.) Theorem 6 tells us that $S(\overline{\psi}) = \infty$ in this case. If ψ is a countable convex combination of pure states then $p\pi_\psi(\mathcal{A})''p$ is a countable direct sum of type I factors and we may refer to Theorem 6 again.

The expression (11) appeared as the definition of entropy in [6]. It is noteworthy that in [9] this formula was generalized to define the entropy of a state with respect to a compact convex subset.

Theorem 6 and Lemma 7 with the additivity of the von Neumann entropy under tensor product gives that

$$S(\varphi \otimes \psi) = S(\varphi) + S(\psi) \tag{12}$$

if φ and ψ are arbitrary states of C^*-algebras. It is quite remarkable that this property does not follow readily from the definition (8).

REFERENCES

[1] ARAKI, H., Recent progress on entropy and relative entropy, *VIIIth Int. Congr. on Math. Phys.* (Marseille, 1986), World Sci. Publishing, Singapore, 1987, 354–365. *MR* **88j**:81004

[2] CSISZÁR, I. and KÖRNER, J., *Information Theory. Coding theorems for discrete memoryless systems*, Probability and Mathematical Statistics, Academic Press, New York–London, 1981. *MR* **84e**:94007

[3] HIAI, F., Minimum index for subfactors and entropy II, *J. Math. Soc. Japan* **43** (1991), 347–379.

[4] JAYNES, E. T., Information theory and statistical mechanics. II, *Phys. Rev.* (2) **108** (1957), 171–190. *MR* **20** #2898

[5] KOSAKI, H., Relative entropy of states: a variational expression, *J. Operator Theory* **16** (1986), 335–348. *MR* **87j**:46110

[6] MANUCEAU, J., NAUDTS, J. and VERBEURE, A., Entropy of normal states, *Comm. Math. Phys.* **27** (1972), 327–338.

[7] NARNHOFER, H. and THIRRING, W., From relative entropy to entropy, *Fizika* **17** (1985), 257–265.

[8] NEUMANN, J. VON, *Mathematische Grundlagen der Quantenmechanik*, Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, Bd. 38, Springer, Berlin, 1932. *Jb.Fortschritte Math.* **58**, 929–930. Dover Publications, New York, N. Y., 1943. *MR* **5**-165

[9] OHYA, M., Entropy transmission in C^*-dynamical systems, *J. Math. Anal. Appl.* **100** (1984), 222–235.

[10] PETZ, D., Properties of quantum entropy, *Quantum probability and applications*, II (Heidelberg, 1984), ed. by L. Accardi and W. von Waldenfels, Lecture Notes in Math., No. 1136, Springer, Berlin–New York, 1985, 428–441. *MR* **87i**:46144

[11] RÉNYI, A., On the foundations of information theory, *Rev. Inst. Internat. Statist.* **33** (1965), 1–14. *MR* **31** #5712

[12] SHANNON, C. E., A mathematical theory of communication, *Bell System Tech. J.* **27** (1948), 379–423. *MR* **10**-133

(Received April 12, 1992)

DEPARTMENT OF INFORMATION SCIENCES
SCIENCE UNIVERSITY OF TOKYO
NODA CITY
CHIBA 278
JAPAN

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 431–443

ON THE LOWER SHADOW IN VARIOUS POSETS

ANDREA RAUSCHE

Abstract

The main aim of this paper is to find new bounds of the lower shadow of families in a poset. At first we consider the Boolean lattice and afterwards the poset of multisets.

1. Introduction

Let n be an integer and let N be an n-element set, e.g. $N=\{1,2,\dots,n\}$. We consider subsets of the set $\binom{N}{k}$ of all k-element subsets of N. The Boolean lattice is the poset of all subsets of N ordered by inclusion. Furthermore, we define the lower shadow $\Delta\mathcal{F}$ of a family $\mathcal{F}\subseteqq\binom{N}{k}$ by $\Delta\mathcal{F}=$ $=\{F\colon \exists G\in\mathcal{F}, F\subset G, |G\setminus F|=1\}$. The following lemma of Sperner [10] gives a simple relationship between $|\mathcal{F}|$ and $|\Delta\mathcal{F}|$:

LEMMA 1. *If $\mathcal{F}\subseteqq\binom{N}{k}$ then $|\Delta\mathcal{F}|\geqq\frac{k}{n-k+1}|\mathcal{F}|$. Equality holds if and only if $\mathcal{F}=\emptyset$ or $\mathcal{F}=\binom{N}{k}$.*

This gives a lower bound for $\frac{|\Delta\mathcal{F}|}{|\mathcal{F}|}$ under the assumption $|\mathcal{F}|>0$. But the inequality is not so good for families $\mathcal{F}$, which sizes are small relative to $\binom{n}{k}$. An important theorem in this field is the Kruskal–Katona Theorem [5], [6]:

THEOREM 1. *Let $\mathcal{F}\subseteqq\binom{N}{k}$ and let*

$$|\mathcal{F}|=\binom{a_k}{k}+\binom{a_{k-1}}{k-1}+\ldots+\binom{a_t}{t}$$

1991 *Mathematics Subject Classification.* Primary 05A05.
Key words and phrases. Shadow, extremal families of subsets.

with $a_k > a_{k-1} > \ldots > a_t \geqq 1$. *Then*

$$|\Delta\mathcal{F}| \geqq \binom{a_k}{k-1} + \binom{a_{k-1}}{k-2} + \ldots + \binom{a_t}{t-1}$$

and this bound is best possible.

Now a useful corollary of Frankl [3] can be proved.

COROLLARY 1. *Let* $\mathcal{F} \subseteqq \binom{N}{k}$ *and let* $|\mathcal{F}| \leqq \binom{m}{k}$ *with* $m \in \mathbb{N}$. *Then*

$$|\Delta\mathcal{F}| \geqq \frac{k}{m-k+1}|\mathcal{F}|, \quad \textit{i.e.}$$

$$\frac{|\Delta\mathcal{F}|}{|\mathcal{F}|} \geqq \frac{\binom{m}{k-1}}{\binom{m}{k}}.$$

Before proving Corollary 1 let us introduce some notations. Consider the (colexicographic) ordering $<_l$ on $\binom{N}{k}$ which is defined by

$$X <_l Y \text{ iff } \max\{v : v \in X - Y\} < \max\{v : v \in Y - X\}.$$

Let $C(m,k)$ denote the first m sets of $\binom{N}{k}$ relative to the ordering $<_l$ and $C\mathcal{F} = C(|\mathcal{F}|, k)$. Then it is not difficult to verify that the Kruskal–Katona Theorem is equivalent to

THEOREM 2. *Let* $\mathcal{F} \subseteqq \binom{N}{k}$. *Then* $|\Delta\mathcal{F}| \geqq |\Delta(C\mathcal{F})|$.

PROOF. $|C(\Delta\mathcal{F})| = |\Delta\mathcal{F}| \geqq |\Delta(C\mathcal{F})|$ due to Theorem 2. With $|\mathcal{F}| \leqq \binom{m}{k}$ we obtain $C\mathcal{F} \subseteqq \binom{M}{k}$, where $M = \{1, 2, \ldots, m\}$. By Lemma 1 we have $|\Delta(C\mathcal{F})| \geqq \frac{k}{m-k+1}|C\mathcal{F}| = \frac{k}{m-k+1}|\mathcal{F}|$. □

2. The main result

Is it possible to improve the inequality of Corollary 1 for certain conditions of $|\mathcal{F}|$? Before we can state and prove our theorem we need the following lemma and some notations:

LEMMA 2. *For all $x, y, z, w \in \mathbb{R}^+$ and $y > w > 0$ in case* (ii) *we have*

$$\text{(i)}\quad (1)\quad \frac{x}{y} \leqq \frac{z}{w} \Leftrightarrow \frac{x}{y} \leqq \frac{x+z}{y+w} \leqq \frac{z}{w}$$

$$(2)\quad \frac{x}{y} \leqq \frac{x+z}{y+w} \Leftrightarrow \frac{x}{y} \leqq \frac{x+z}{y+w} \leqq \frac{z}{w}$$

$$(3)\quad \frac{x+z}{y+w} \leqq \frac{z}{w} \Leftrightarrow \frac{x}{y} \leqq \frac{x+z}{y+w} \leqq \frac{z}{w}$$

$$\text{(ii)}\quad \frac{x}{y} \leqq \frac{x-z}{y-w} \Leftrightarrow \frac{x}{y} \geqq \frac{z}{w}.$$

Now we use $k, m, t \in \mathbb{N}^+$ with k fixed and

$$|\mathcal{F}| = m := \sum_{i=t}^{k} \binom{x_i}{i}$$

with $x_k > x_{k-1} > \ldots > x_t \geqq t \geqq 1$ and all $x_i \in \mathbb{N}$. By Theorem 1 we obtain $|\Delta\mathcal{F}| \geqq \Delta m := \sum_{i=t}^{k} \binom{x_i}{i-1}$. The exact values of x_i can be seen in the following.

Now we are able to state our main result.

THEOREM 3. *Let $\mathcal{F} \subseteqq \binom{N}{k}$ satisfy*

$$|\mathcal{F}| = \binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-1} + \binom{b_{v+1}}{k-v-1}$$

with $v < k-1$ and $a > b_1 > b_2 > \ldots > b_v > b_{v+1}$. Suppose that

$$|\mathcal{F}| \leqq \binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i}$$

implies

$$\frac{|\Delta\mathcal{F}|}{|\mathcal{F}|} \geqq \frac{\binom{a}{k-1} + \sum_{i=1}^{v} \binom{b_i}{k-i-1}}{\binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i}}.$$

Then

(i) *for*

$$\frac{\binom{b_{v+1}}{k-v-2}}{\binom{b_{v+1}}{k-v-1}} \geqq \frac{\binom{a}{k-1}+\sum_{i=1}^{v}\binom{b_i}{k-i-1}}{\binom{a}{k}+\sum_{i=1}^{v}\binom{b_i}{k-i}}$$

we have

$$\frac{|\Delta\mathcal{F}|}{|\mathcal{F}|} \geqq \frac{\binom{a}{k-1}+\sum_{i=1}^{v}\binom{b_i}{k-i-1}}{\binom{a}{k}+\sum_{i=1}^{v}\binom{b_i}{k-1}};$$

(ii) *for*

$$\frac{\binom{b_{v+1}}{k-v-2}}{\binom{b_{v+1}}{k-v-1}} \leqq \frac{\binom{a}{k-1}+\sum_{i=1}^{v}\binom{b_i}{k-i-1}}{\binom{a}{k}+\sum_{i=1}^{v}\binom{b_i}{k-i}},$$

we have

$$\frac{|\Delta\mathcal{F}|}{|\mathcal{F}|} \geqq \frac{\binom{a}{k-1}+\sum_{i=1}^{v}\binom{b_i}{k-i-1}+\binom{b_{v+1}}{k-v-2}}{\binom{a}{k}+\sum_{i=1}^{v}\binom{b_i}{k-i}+\binom{b_{v+1}}{k-v-1}}.$$

REMARK 1. If

$$\frac{\binom{b_{v+1}}{k-v-2}}{\binom{b_{v+1}}{k-v-1}} = \frac{\binom{a}{k-1}+\sum_{i=1}^{v}\binom{b_i}{k-i-1}}{\binom{a}{k}+\sum_{i=1}^{v}\binom{b_i}{k-i}},$$

then by Lemma 2 we obtain the same results as in Theorem 3(i) and as in Theorem 3(ii).

The case $v=0$ means $|\mathcal{F}| \leqq \binom{a}{k}+\binom{b}{k-1}$ what was discussed by the author in [8].

3. Proof of Theorem 3

(i) The assumptions imply

$$\frac{|\Delta\mathcal{F}|}{|\mathcal{F}|} \geqq \frac{\binom{a}{k-1} + \sum_{i=1}^{v} \binom{b_i}{k-i-1}}{\binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i}}$$

for $m \leqq \binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i}$. Now let

$$\binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i} + 1 \leqq m \leqq \binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i} + \binom{b_{v+1}}{k-v-1}$$

and we have to investigate

$$\frac{\Delta m}{m} = \frac{\binom{a}{k-1} + \sum_{i=1}^{v} \binom{b_i}{k-i-1} + \sum_{i=v+1}^{k-t} \binom{c_i}{k-i-1}}{\binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i} + \sum_{i=v+1}^{k-t} \binom{c_i}{k-i}}$$

with $c_{v+1} < b_{v+1}$.

By Theorem 1 and Corollary 1 we know

$$\frac{\sum_{i=v+1}^{k-t} \binom{c_i}{k-i-1}}{\sum_{i=v+1}^{k-t} \binom{c_i}{k-i}} \geqq \frac{\binom{b_{v+1}}{k-v-2}}{\binom{b_{v+1}}{k-v-1}}$$

and with our assumption we obtain

$$\frac{\sum_{i=v+1}^{k-t} \binom{c_i}{k-i-1}}{\sum_{i=v+1}^{k-t} \binom{c_i}{k-i}} \geqq \frac{\binom{a}{k-1} + \sum_{i=1}^{v} \binom{b_i}{k-i-1}}{\binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i}}.$$

Then from Lemma 2 we infer

$$\frac{\Delta m}{m} \geqq \frac{\binom{a}{k-1} + \sum_{i=1}^{v} \binom{b_i}{k-i-1}}{\binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i}}.$$

(ii) Since

$$\frac{|\Delta\mathcal{F}|}{|\mathcal{F}|} \geqq \frac{\binom{a}{k-1} + \sum_{i=1}^{v}\binom{b_i}{k-i-1} + \binom{b_{v+1}}{k-v-2}}{\binom{a}{k} + \sum_{i=1}^{v}\binom{b_i}{k-i} + \binom{b_{v+1}}{k-v-1}}$$

for

$$m \leqq \binom{a}{k} + \sum_{i=1}^{v}\binom{b_i}{k-i}$$

(see assumptions and Lemma 2) we consider the interval

$$\binom{a}{k} + \sum_{i=1}^{v}\binom{b_i}{k-i} + 1 \leqq m \leqq \binom{a}{k} + \sum_{i=1}^{v}\binom{b_i}{k-i} + \binom{b_{v+1}}{k-v-1}.$$

According to the further proof it is necessary to decompose this interval. Therefore let

$$m \in \left(\binom{a}{k} + \sum_{i=1}^{v}\binom{b_i}{k-i} + \binom{b_{v+1}-1}{k-v-1}, \binom{a}{k} + \sum_{i=1}^{v}\binom{b_i}{k-i} + \binom{b_{v+1}}{k-v-1}\right)$$

and

$$\frac{\Delta m}{m} = \frac{\binom{a}{k-1} + \sum_{i=1}^{v}\binom{b_i}{k-i-1} +}{\binom{a}{k} + \sum_{i=1}^{v}\binom{b_i}{k-i} +} \cdots$$

$$\cdots \frac{+\binom{b_{v+1}-1}{k-v-2} + \sum_{i=v+2}^{k-t}\binom{b_i}{k-i-1}}{+\binom{b_{v+1}-1}{k-v-1} + \sum_{i=v+2}^{k-t}\binom{b_i}{k-i}}.$$

By Corollary 1 and Theorem 1 we obtain

$$\frac{\Delta m'}{m'} := \frac{\binom{b_{v+1}-1}{k-v-2} + \sum_{i=v+2}^{k-t}\binom{b_i}{k-i-1}}{\binom{b_{v+1}-1}{k-v-1} + \sum_{i=v+2}^{k-t}\binom{b_i}{k-i}} \geqq \frac{\binom{b_{v+1}}{k-v-2}}{\binom{b_{v+1}}{k-v-1}}.$$

Hence

$$-\Delta m'\binom{b_{v+1}}{k-v-1} \leqq -m'\binom{b_{v+1}}{k-v-2}$$

and we get

$$\frac{\binom{b_{v+1}}{k-v-2} - \Delta m'}{\binom{b_{v+1}}{k-v-1} - m'} \leqq \frac{\binom{b_{v+1}}{k-v-2}}{\binom{b_{v+1}}{k-v-1}} \leqq \frac{\binom{a}{k-1} + \sum_{i=1}^{v}\binom{b_i}{k-i-1}}{\binom{a}{k} + \sum_{i=1}^{v}\binom{b_i}{k-i}}$$

due to the assumptions. (Notice that $\binom{b_{v+1}}{k-v-1} - m' > 0$ because of the definition of m'.) Then, using Lemma 2, we obtain

$$\frac{\binom{b_{v+1}}{k-v-2} - \Delta m'}{\binom{b_{v+1}}{k-v-1} - m'} \leqq \frac{\binom{a}{k-1} + \sum_{i=1}^{v}\binom{b_i}{k-i-1} + \binom{b_{v+1}}{k-v-2}}{\binom{a}{k} + \sum_{i=1}^{v}\binom{b_i}{k-i} + \binom{b_{v+1}}{k-v-1}}.$$

This holds by Lemma 2 if and only if

$$\frac{\Delta m}{m} = \frac{\binom{a}{k-1} + \sum_{i=1}^{v}\binom{b_i}{k-i-1} + \binom{b_{v+1}}{k-v-2} - \binom{b_{v+1}}{k-v-2} + \Delta m'}{\binom{a}{k} + \sum_{i=1}^{v}\binom{b_i}{k-i} + \binom{b_{v+1}}{k-v-1} - \binom{b_{v+1}}{k-v-1} + m'}$$
$$\geqq \frac{\binom{a}{k-1} + \sum_{i=1}^{v}\binom{b_i}{k-i-1} + \binom{b_{v+1}}{k-v-2}}{\binom{a}{k} + \sum_{i=1}^{v}\binom{b_i}{k-i} + \binom{b_{v+1}}{k-v-1}}.$$

Let $x' \in \mathbb{N}$ satisfy

$$\frac{\binom{x'}{k-v-2}}{\binom{x'}{k-v-1}} \leqq \frac{\binom{a}{k-1} + \sum_{i=1}^{v}\binom{b_i}{k-i-1}}{\binom{a}{k} + \sum_{i=1}^{v}\binom{b_i}{k-i}} \leqq \frac{\binom{x'-1}{k-v-2}}{\binom{x'-1}{k-v-1}}.$$

We have to prove that in every interval

$$m \in \left(\binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i} + \binom{x'+i-1}{k-v-1}, \binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i} + \binom{x'+i}{k-v-1} \right];$$

$i = 0, 1, \ldots, b_{v+1} - x'$; the statement

$$\frac{\Delta m}{m} \geqq \frac{\binom{a}{k-1} + \sum_{i=1}^{v} \binom{b_i}{k-i-1} + \binom{x'+i}{k-v-2}}{\binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i} + \binom{x'+i}{k-v-1}}$$

is satisfied. The interval

$$\binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i} + 1 \leqq m \leqq \binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i} + \binom{b_{v+1}}{k-v-1}$$

can be partitioned into:

$$m \in \left(\binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i}, \binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i} + \binom{x'-1}{k-v-1} \right] \cup$$

$$\bigcup_{i=0}^{b_{v+1}-x'} \left(\binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i} + \binom{x'+i-1}{k-v-1}, \binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i} + \binom{x'+i}{k-v-1} \right].$$

Therefore we have previously investigated the intervals of the second part and we made sure that

$$\frac{|\Delta \mathcal{F}|}{|\mathcal{F}|} \geqq \frac{\binom{a}{k-1} + \sum_{i=1}^{v} \binom{b_i}{k-i-1} + \binom{b_{v+1}}{k-v-2}}{\binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i} + \binom{b_{v+1}}{k-v-1}}$$

is true for the first part of the interval.

Now we have only to find out: What is the relationship between the minimums of the several intervals of the second part? In order to show that the desired minimum lies at the right boundary we prove

$$
\frac{\binom{a}{k-1}+\sum_{i=1}^{v}\binom{b_i}{k-i-1}+\binom{x'+y}{k-v-2}}{\binom{a}{k}+\sum_{i=1}^{v}\binom{b_i}{k-i}+\binom{x'+y}{k-v-1}} \geqq
\tag{1}
$$

$$
\geqq \frac{\binom{a}{k-1}+\sum_{i=1}^{v}\binom{b_i}{k-i-1}+\binom{x'+y+1}{k-v-2}}{\binom{a}{k}+\sum_{i=1}^{v}\binom{b_i}{k-i}+\binom{x'+y+1}{k-v-1}}
$$

for arbitrary y; $0 \leqq y \leqq b_{v+1}-x'-1$; and with our assumption

$$
\frac{\binom{a}{k-1}+\sum_{i=1}^{v}\binom{b_i}{k-i-1}}{\binom{a}{k}+\sum_{i=1}^{v}\binom{b_i}{k-i}} \geqq \frac{\binom{u}{k-v-2}}{\binom{u}{k-v-1}};
$$

$x' \leqq u \leqq b_{v+1}$. If we transform (1) it is easy to see that

$$
\begin{aligned}
&\binom{a}{k-1}\left[\binom{x'+y+1}{k-v-1}-\binom{x'+y}{k-v-1}\right] \\
&+\binom{x'+y}{k-v-2}\left[\binom{a}{k}+\sum_{i=1}^{v}\binom{b_i}{k-i}+\binom{x'+y+1}{k-v-1}\right] \\
&\qquad\qquad+\binom{x'+y+1}{k-v-1}\sum_{i=1}^{v}\binom{b_i}{k-i-1} \\
&\geqq \binom{x'+y+1}{k-v-2}\left[\binom{a}{k}+\sum_{i=1}^{v}\binom{b_i}{k-i}+\binom{x'+y}{k-v-1}\right]
\end{aligned}
$$

must hold. In other words, the inequality

$$
\frac{\binom{x'+y}{k-v-3}}{\binom{x'+y}{k-v-2}} \leqq \frac{\binom{a}{k-1}+\sum_{i=1}^{v}\binom{b_i}{k-i-1}+\binom{x'+y}{k-v-2}}{\binom{a}{k}+\sum_{i=1}^{v}\binom{b_i}{k-i}+\binom{x'+y}{k-v-1}}
\tag{2}
$$

has to be satisfied.

In order to complete our proof we need a family $\mathcal{A} \subseteq \binom{N}{k-v-1}$ with $|\mathcal{A}| = \binom{x'+y}{k-v-1}$. Then, using Theorem 1, $|\Delta\mathcal{A}| \geqq \binom{x'+y}{k-v-2}$ follows. Otherwise Corollary 1 implies

$$\frac{|\Delta\mathcal{A}|}{|\mathcal{A}|} \geqq \frac{\binom{x'+y+1}{k-v-2}}{\binom{x'+y+1}{k-v-1}}$$

since $\binom{x'+y}{k-v-1} < \binom{x'+y+1}{k-v-1}$. We get

$$\frac{\binom{x'+y}{k-v-2}}{\binom{x'+y}{k-v-1}} \geqq \frac{\binom{x'+y+1}{k-v-2}}{\binom{x'+y+1}{k-v-1}} = \frac{\binom{x'+y}{k-v-2} + \binom{x'+y}{k-v-3}}{\binom{x'+y}{k-v-1} + \binom{x'+y}{k-v-2}}$$

since the bound of Theorem 1 is best possible. Hence we obtain

$$\frac{\binom{x'+y}{k-v-2}}{\binom{x'+y}{k-v-1}} \geqq \frac{\binom{x'+y}{k-v-3}}{\binom{x'+y}{k-v-2}} \tag{3}$$

by Lemma 2. On the other hand we have assumed

$$\frac{\binom{x'+y}{k-v-2}}{\binom{x'+y}{k-v-1}} \leqq \frac{\binom{a}{k-1} + \sum_{i=1}^{v} \binom{b_i}{k-i-1}}{\binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i}}$$

and

$$\frac{\binom{x'+y}{k-v-2}}{\binom{x'+y}{k-v-1}} \leqq \frac{\binom{a}{k-1} + \sum_{i=1}^{v} \binom{b_i}{k-i-1} + \binom{x'+y}{k-v-2}}{\binom{a}{k} + \sum_{i=1}^{v} \binom{b_i}{k-i} + \binom{x'+y}{k-v-1}}$$

follows by Lemma 2.

With the help of the last inequality and (3) we have proved (2) and the proof is complete. □

4. Further results

Taking into account that

$$\sum_{i=0}^{t} \binom{a-i}{k-i} = \binom{a+1}{k} - \binom{a-t}{k-t-1}$$

and

$$\sum_{i=0}^{t} \binom{a-i}{k-i-1} = \binom{a+1}{k-1} - \binom{a-t}{k-t-2},$$

we can reformulate our theorem in the following manner:

COROLLARY 2. *Let* $\mathcal{F} \subseteq \binom{N}{k}$ *with* $|\mathcal{F}| \leqq \binom{a+1}{k} - \binom{a-t}{k-t-1}$ *and* $t < k-1 < k \leqq a$. *Then*

$$\frac{|\Delta\mathcal{F}|}{|\mathcal{F}|} \geqq \frac{\binom{a+1}{k-1} - \binom{a-t}{k-t-2}}{\binom{a+1}{k} - \binom{a-t}{k-t-1}}.$$

The proof is given in [9].

Using this corollary and replacing $a+1 \Rightarrow n$ and $t+1 \Rightarrow m$, we obtain a very important result of Griggs [4]:

THEOREM 4. *Let* $\mathcal{F} \subseteq \binom{N}{k}$ *with* $1 \leqq |\mathcal{F}| \leqq \binom{n}{k} - \binom{n-m}{k-m}$, $m \in \mathbb{N}$ *and* $m \leqq k-1$. *Then*

$$\frac{|\Delta\mathcal{F}|}{|\mathcal{F}|} \geqq \frac{\binom{n}{k-1} - \binom{n-m}{k-m-1}}{\binom{n}{k} - \binom{n-m}{k-m}}.$$

5. The multisets $S(k_n, k_{n-1}, \ldots, k_1)$

Let $S := S(k_n, k_{n-1}, \ldots, k_1)$, $0 < k_n \leqq k_{n-1} \leqq \ldots \leqq k_1$, be the set of all vectors $\mathbf{x} = (x_n, x_{n-1}, \ldots, x_1)$ with coordinates satisfying $0 \leqq x_i \leqq k_i$ for all i, $i = 1, \ldots, n$. Let $\begin{bmatrix} j \\ i \end{bmatrix}$ denote the coefficient of x^i in $\prod_{r=1}^{j} (1+x+x^2+\ldots+x^{k_r})$. Furthermore, $N^i(S)$ consists of all vectors of S with coordinate sum i.

For a given vector $\mathbf{x}$ we define $\Delta(\mathbf{x})$ by

$$\Delta(\mathbf{x}) := \{(x_n - 1, x_{n-1}, \dots, x_1), (x_n, x_{n-1} - 1, x_{n-2}, \dots, x_1), \dots \\ \dots, (x_n, \dots, x_2, x_1 - 1)\} \cap S.$$

Of course, the lower shadow of a family $\mathcal{F}$ is the union of the lower shadows of the elements of $\mathcal{F}$.

In order to answer the question how to calculate or estimate $|\Delta\mathcal{F}|$ for fixed $|\mathcal{F}|$ over all $\mathcal{F} \subseteq N^k(S)$, Clements [2] developed the following method. We will present it here without proof.

THEOREM 5. *Let i, m and $k_n \leqq k_{n-1} \leqq \dots \leqq k_1$ be given positive integers satisfying $1 \leqq m \leqq \begin{bmatrix} n \\ i \end{bmatrix}$ and $1 \leqq i \leqq k_n + \dots + k_1$. Then there exist uniqe integers $m(i), m(i-1), \dots, m(t)$, $i \geqq t > 0$, such that*

(1) $$m = \begin{bmatrix} m(i) \\ i \end{bmatrix} + \begin{bmatrix} m(i-1) \\ i-1 \end{bmatrix} + \dots + \begin{bmatrix} m(t) \\ t \end{bmatrix}.$$

We will refer to this as the i-representation of m.

(2) $$m(i) \geqq m(i-1) \geqq \dots \geqq m(t) \quad \textit{and} \quad \begin{bmatrix} m(i) \\ i \end{bmatrix} > 0.$$

(3) *If $\mathcal{F} \subseteq N^i(S)$ with $|\mathcal{F}| = m$, then*

$$|\Delta\mathcal{F}| \geqq \begin{bmatrix} m(i) \\ i-1 \end{bmatrix} + \begin{bmatrix} m(i-1) \\ i-2 \end{bmatrix} + \dots + \begin{bmatrix} m(t) \\ t-1 \end{bmatrix}.$$

The following lemma is very important.

LEMMA 3. *Let $\mathcal{F} \subseteq N^i(S)$ with $0 < |\mathcal{F}| \leqq \begin{bmatrix} m(i) \\ i \end{bmatrix}$, then*

$$\frac{|\Delta\mathcal{F}|}{|\mathcal{F}|} \geqq \frac{\begin{bmatrix} m(i) \\ i-1 \end{bmatrix}}{\begin{bmatrix} m(i) \\ i \end{bmatrix}}.$$

Taking into account that Theorem 5 is a generalization of the Kruskal–Katona Theorem we are able to formulate the next theorem:

THEOREM 6. *Let $\mathcal{F} \subseteq N^k(S)$ with $0 < |\mathcal{F}| \leqq \begin{bmatrix} a \\ k \end{bmatrix} + \begin{bmatrix} b \\ k-1 \end{bmatrix}$, $n > a \geqq b \geqq 1$ and $\begin{bmatrix} b \\ k-1 \end{bmatrix}$. Then*

(i) $$\frac{\begin{bmatrix} b \\ k-2 \end{bmatrix}}{\begin{bmatrix} b \\ k-1 \end{bmatrix}} \geqq \frac{\begin{bmatrix} a \\ k-1 \end{bmatrix}}{\begin{bmatrix} a \\ k \end{bmatrix}} \quad \textit{implies} \quad \frac{|\Delta\mathcal{F}|}{|\mathcal{F}|} \geqq \frac{\begin{bmatrix} a \\ k-1 \end{bmatrix}}{\begin{bmatrix} a \\ k \end{bmatrix}},$$

$$\text{(ii)} \qquad \frac{\begin{bmatrix} b \\ k-2 \end{bmatrix}}{\begin{bmatrix} b \\ k-1 \end{bmatrix}} \leqq \frac{\begin{bmatrix} a \\ k-1 \end{bmatrix}}{\begin{bmatrix} a \\ k \end{bmatrix}} \ \textit{implies} \ \frac{|\Delta\mathcal{F}|}{|\mathcal{F}|} \geqq \frac{\begin{bmatrix} a \\ k-1 \end{bmatrix} + \begin{bmatrix} b \\ k-2 \end{bmatrix}}{\begin{bmatrix} a \\ k \end{bmatrix} + \begin{bmatrix} b \\ k-1 \end{bmatrix}}.$$

Having replaced v by 0, the proof is similar to that of Theorem 3.

6. Concluding remarks

These results suggest that for all posets of the so-called Kruskal–Katona type a theorem analogous to Theorem 3 may hold. In [9] we found results for the poset of Leeb [7] and formulated a conjecture for a poset of Bezrukov and Gronau [1].

REFERENCES

[1] Bezrukov, S. L. and Gronau, H.-D. O. F., A Kruskal–Katona type theorem (preprint).

[2] Clements, G. F., A generalization of the Kruskal–Katona theorem, *J. Combin. Theory. Ser. A* **37** (1984), 91–97. *MR* **85i**:05014

[3] Frankl, P., On Sperner families satisfying an additional condition, *J. Combinatorial Theory Ser. A* **20** (1976), 1–11. *MR* **53** #2693

[4] Griggs, J. R., Collections of subsets with the Sperner property, *Trans. Amer. Math. Soc.* **269** (1982), 575–591. *MR* **83d**:05003

[5] Katona, G. O. H., A theorem of finite sets, *Theory of graphs* (Proc. Colloq., Tihany, 1966), Academic Press, New York, 1968, 187–207. *MR* **45** #76

[6] Kruskal, J. B., The number of simplices in a complex, *Mathematical optimization techniques*, Univ. of California Press, Berkeley, Calif., 1963, 251–278. *MR* **27** #4771

[7] Leeb, K., Salami-Taktik beim Quader-Packen, Arbeitsber. Inst. Math. Masch. Datenverarb. (Inform.) **11** (1978), no. 5, 1–15. *MR* **83d**:52014

[8] Rausche, A., On an inequality of Sperner, *Topics in combinatorics and graph theory* (Oberwolfach, 1990), Physica-Verlag, Heidelberg, 1990, 569–576. *MR* **92b**: 05076

[9] Rausche, A., Sperner-Schattenabschätzungen in Halbordnungen, Dissertation, Greifswald, 1991.

[10] Sperner, E., Ein Satz über Untermengen einer endlichen Menge, *Math. Z.* **27** (1928), 544–548. *Jb. Fortschritte Math.* **54**, 90–91

(Received October 23, 1991)

MATHEMATISCHES INSTITUT DER
ERNST-MORITZ-ARNDT-UNIVERSITÄT
FRIEDRICH-LUDWIG-JAHN-STRASSE 15A
GREIFSWALD
FEDERAL REPUBLIC OF GERMANY

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 445–453

A GENERALIZATION OF MIXED AREAS AND THE GREATEST EGYPTIAN PYRAMID

G. D. CHAKERIAN and M. A. MNATSAKANIAN

Respectfully dedicated to Professor László Fejes Tóth on the occasion of his 80th birthday

1. The volume of a prismoid. An Egyptian formula for the volume V of a truncated square pyramid, dubbed the "Greatest Egyptian Pyramid" by Eric Temple Bell, is found in the Moscow Papyrus (circa 1900 B.C.), and is presented there as a special numerical case of the formula

$$V = \frac{h}{3}(a^2 + ab + b^2), \tag{1.1}$$

where a and b are the sidelengths of the bottom and top square bases, and h is the height, i.e., the distance between the parallel planes containing the square faces. For more historical details, see van der Waerden [7]. Rewriting this formula in terms of the areas $A_0 = a^2$ and $A_1 = b^2$ of the square faces gives

$$V = \frac{h}{3}(A_0 + \sqrt{A_0 A_1} + A_1). \tag{1.2}$$

In fact, it is known that (1.2) is valid for the volume of the solid resulting from the truncation of an arbitrary cone by a plane parallel to the base, where A_0 and A_1 are the areas of the bottom and top bases.

More generally, consider parallel planes π_0 and π_1 in three-dimensional Euclidean space E^3, with compact plane convex sets $K_0 \subset \pi_0$ and $K_1 \subset \pi_1$. Then the convex hull of $K_0 \cup K_1$ is a convex set K which we shall call a "generalized prismoid". In case K_0 and K_1 are convex polygons, K is a convex polytope usually called a prismoid (or a prismatoid). Taking a coordinate system in E^3 with z-axis orthogonal to π_0 and π_1, with π_0 the plane $z=0$ and π_1 the plane $z=h$, where h is the height of K, one sees that the cross-section of K by the parallel plane at height z above π_0, $0 \leq z \leq h$, is the Minkowski convex combination $(1-\frac{z}{h})K_0 + (\frac{z}{h})K_1$. The area of this cross-section is

$$A\left(\left(1-\frac{z}{h}\right)K_0+\left(\frac{z}{h}\right)K_1\right) = \left(1-\frac{z}{h}\right)^2 A_0+2\left(\frac{z}{h}\right)\left(1-\frac{z}{h}\right)A_{01}+\left(\frac{z}{h}\right)^2 A_1, \tag{1.3}$$

where $A_0 = A(K_0)$ and $A_1 = A(K_1)$ are the areas of K_0 and K_1, respectively, and $A_{01} = A(K_0, K_1)$ is the mixed area as introduced by Minkowski (circa

1991 *Mathematics Subject Classification.* Primary 52A38, 53A04; Secondary 52A39.
Key words and phrases. Volume of a prismoid, mixed areas, linking vector.

0081–6906/96/$ 5.00

A.D. 1900). It is to be noted that $A(K_0, K_1)$ is taken here to mean the mixed area of K_0 with any translate of K_1 lying in the same plane and is invariant under translations in E^3 of either K_0 or K_1. Integrating the cross-sectional area (1.3) over the interval $[0, h]$ yields the volume $V(K)$,

$$V(K) = \frac{h}{3}(A_0 + A_{01} + A_1), \tag{1.4}$$

which expresses the volume of K in terms of K_0, K_1, and the height.

If K is a truncated cone, so K_0 and K_1 are homothetic (i.e., K_0 is a translate of λK_1 for some $\lambda > 0$), then it is easily checked that $A_{01} = \sqrt{A_0 A_1}$, and (1.4) reduces to the classical case (1.2).

Minkowski's inequality is

$$A_{01} \geqq \sqrt{A_0 A_1}, \tag{1.5}$$

with equality if and only if K_0 and K_1 are homothetic. From this we see that the expression (1.2) would be an underestimate for the volume of a generalized prismoid except in the case of a truncated cone.

As an interesting historical aside on (1.4), Heron in Book II of the Metrica (see Heath [4], p. 333) gave a formula equivalent to

$$V = \frac{h}{3}\left(ab + \frac{bc + ad}{2} + cd\right) \tag{1.6}$$

for the case where K_0 is a rectangle with sides of lengths a, b and K_1 a rectangle with sides of lengths c, d parallel to those of K_0. Note that $(bc + ad)/2$ is indeed the mixed area of these two rectangles.

We conclude this section with another expression for *the volume of a generalized prismoid, which together with (1.4) will be extended to more general solids* in later sections. Let M be the area of the cross-section of the generalized prismoid K midway between the planes π_0 and π_1. Then we have

$$M = A\left(\frac{K_0 + K_1}{2}\right) = \frac{1}{4}(A_0 + 2A_{01} + A_1). \tag{1.7}$$

This gives $A_{01} = 2M - (A_0 + A_1)/2$ and substitution in (1.4) yields,

$$V(K) = \frac{h}{6}(A_0 + 4M + A_1). \tag{1.8}$$

This is the well-known formula for the convex prismoidal volumes (see Pólya [6], Vol. I, pp. 110–113). In fact, it also follows directly from Simpson's approximation rule for integrals, which in this case gives the exact result, since the volume of K is obtained by integrating a *quadratic* function of z. We shall see that the generalization to *nonconvex* prismoidal volumes will be valid for the same reason.

2. A vector invariant for a closed space curve. Let C be a closed curve in E^3 parametrized as $x = x(t)$, $0 \leqq t \leqq T$. In the following we shall assume, unless otherwise stated, that C is a smooth curve, and define the *vector invariant* $I(C)$ by

$$(2.1) \qquad I(C) = \frac{1}{2} \oint_C [x, dx] = \frac{1}{2} \int_0^T [x(t), \dot{x}(t)] dt,$$

where $[u, v]$ denotes the usual cross product (or outer product) of vectors u and v in E^3. Note that, except for sign, $I(C)$ is independent of the parametrization of C as a closed curve.

The vector $I(C)$ admits the following *geometric interpretation:* Its magnitude is the maximum of the algebraic areas of the orthogonal projections of C on planes, and its direction is normal to the plane that yields this maximum. In fact, it is not difficult to show that for any unit vector u, the quantity $u \cdot I(C)$ is the algebraic area determined by the orthogonal projection of C into a plane orthogonal to u. That is, if C^* is the orthogonal projection of C into a plane E orthogonal to u, with C^* given the orientation induced by C and E the orientation consistent with the normal vector u, then

$$(2.2) \qquad u \cdot I(C) = \iint_E w(C^*, p) dA(p),$$

where $w(C^*, p)$ is the winding number of $p \in E$ with respect to C^*, and $dA(p)$ is the area element in E (the set of points $p \in E$ for which $w(C^*, p)$ is undefined is C^* and has measure zero).

An application of the theorem of Stokes shows that if C is the boundary of any smooth orientable surface S, and n is a smoothly varying unit normal vector field on S, then

$$(2.3) \qquad I(C) = \pm \iint_S n dS,$$

where dS denotes the surface area element of S. (In case C is a simple closed curve lying in a plane, $I(C)$ is a vector orthogonal to the plane with magnitude equal to the area enclosed by C.) Thus we see, by dotting both sides of (2.3) with a unit vector u, that the algebraic area in the preceding geometric interpretation is the algebraic area of the projection of any orientable surface S having C as its boundary. In particular, the definition (2.1) relates directly to the case where S is a cone with vertex at the origin.

The geometric interpretation of $I(C)$ in terms of algebraic areas of projections leads to the following unexpected conclusion. Position an arbitrary closed space curve so that its projection into a horizontal plane provides the maximum algebraic area. Then all of its projections on vertical planes have vanishing algebraic area.

One can give *another* interesting *geometric interpretation* in terms of linking numbers. Given any oriented line G in E^3 with direction u, if E is any plane orthogonal to u and $G \cap E = \{p\}$, then the winding number $w(C^*, p)$ is the linking number of G and C, assuming $G \cap C = \emptyset$. We now define the *vectorial linking number*, or the *linking vector*, $\lambda(C, G)$ to be $w(C^*, p)u$. From (2.2) we then have

$$u(u \cdot I(C)) = \iint_E uw(C^*, p)dA(p) = \iint_E \lambda(C, G)dA(p), \tag{2.4}$$

where the oriented lines G in the last integral are those having orientation u orthogonal to E, intersecting E in the points p. If S^2 is the standard unit sphere, we obtain from (2.4),

$$\int_{S^2} \{u(u \cdot I(C))\}d\sigma(u) = \int_{S^2} \left\{\iint_E \lambda(C, G)dA(p)\right\}d\sigma(u), \tag{2.5}$$

where $d\sigma(u)$ is the area element of S^2. In (2.5), one easily verifies that the integral on the left-hand side is $\omega_3 I(C)$ (ω_3 is the volume of the unit ball). The motion invariant integral-geometric density for oriented lines in E^3 is $dG = dAd\sigma(u)$, where u is the direction of G, $d\sigma(u)$ the area element of S^2 at point u, and dA the area element on a plane orthogonal to u at the point where the plane intersects G. Thus the integral on the right-hand side of (2.5) is the integral of $\lambda(C, G)$ with respect to the density for lines, and we have

$$\frac{4\pi}{3}I(C) = \int \lambda(C, G)dG, \tag{2.6}$$

where the integral is over all oriented lines in E^3.

Observe that in the definition of $\lambda(C, G)$, when the orientation of G is reversed, so is the orientation of the plane into which C is projected to compute the winding number; thus the linking vector is in fact unchanged. This explains why the integral in (2.6) does not suffer the drawback of being always zero, as in the case of the similar integral for linking numbers of space curves mentioned by Blaschke [1, p. 120].

It should be mentioned that an analogous vector invariant can be defined in a natural way for a compact $(n-2)$-dimensional manifold embedded in E^n, namely by integrating the generalized cross-product $[x, dx, \dots, dx]/(n-1)!$ over the manifold. There are $(n-2)$ dx's in the product, which is calculated using exterior multiplication (see Flanders [3] for another application).

All the properties discussed in this section generalize in a straightforward fashion, including the analogues of (2.2), (2.3), (2.6), and a geometric interpretation in terms of the algebraic $(n-1)$-dimensional volumes of the projections into hyperplanes.

3. A mixed vector invariant. We now consider two smooth closed curves C_0 and C_1 in E^3 with parametrizations $x_0(t)$ and $x_1(t)$, such that both C_0 and C_1 are traversed once as t varies from 0 to T. We define a *mixed vector invariant* $I(C_0, C_1)$ by

$$I(C_0, C_1) = \frac{1}{2} \oint_{C_1} [x_0, dx_1] = \frac{1}{2} \int_0^T [x_0(t), \dot{x}_1(t)] dt. \tag{3.1}$$

The derivative of $[x_0, x_1]$ is $[x_0, \dot{x}_1] - [x_1, \dot{x}_0]$ and integration of this gives the symmetry relation

$$I(C_0, C_1) = I(C_1, C_0). \tag{3.2}$$

One must be careful to observe that the notation is deceptive since, in contrast to the invariant $I(C)$, this mixed invariant is strongly dependent on the parametrizations of C_0 and C_1.

In case C_0 and C_1 are plane curves lying in a common plane, $\pm\|I(C_0, C_1)\|$ is the invariant defined as in Chakerian and Goodey [2]. In particular, if C_0 and C_1 are smooth strictly convex plane curves parametrized in such a way that the outward normal vectors at corresponding points are in the same direction, then $\|I(C_0, C_1)\| = A(K_0, K_1)$, the mixed area of the convex sets K_0 and K_1 bounded by C_0 and C_1.

An easy *bilinearity property* of $I(C_0, C_1)$, analogous to that of mixed areas, is as follows. Let C_0 and C_1 be any closed space curves with parametrizations $x_0(t)$ and $x_1(t)$, $0 \leqq t \leqq T$, as designated at the beginning of this section. Let α_0 and α_1 be any real numbers and C the closed curve given by $x = x(t) = \alpha_0 x_0(t) + \alpha_1 x_1(t)$. The bilinearity of the cross-product then gives

$$I(C) = \frac{1}{2} \oint_C [x, dx] = \alpha_0^2 I(C_0) + 2\alpha_0 \alpha_1 I(C_0, C_1) + \alpha_1^2 I(C_1). \tag{3.3}$$

The invariant $I(C_0, C_1)$ also has a *projection property* similar to that of $I(C)$ discussed in Section 2. If u is any unit vector and C_0^*, C_1^* are the orthogonal projections of C_0, C_1 into the plane E through the origin orthogonal to u, then

$$u \cdot I(C_0, C_1) = \pm\|I(C_0^*, C_1^*)\|, \tag{3.4}$$

where $I(C_0^*, C_1^*)$ is computed using the parametrizations $x_0^*(t)$ and $x_1^*(t)$ of C_0^* and C_1^* induced by the projection. Indeed, one proves (3.4) by writing

$x_0(t) = x_0^*(t) + l_0(t)u$ and $x_1(t) = x_1^*(t) + l_1(t)u$, $0 \leqq t \leqq T$, from which it immediately follows that $u \cdot [x_0, dx_1] = u \cdot [x_0^*, dx_1^*]$, and we have

$$\begin{aligned} u \cdot I(C_0, C_1) &= u \cdot \oint_{C_1} [x_0, dx_1] = \oint_{C_1} u \cdot [x_0, dx_1] = \\ &= \oint_{C_1^*} u \cdot [x_0^*, dx_1^*] = u \cdot \int_{C_1^*} [x_0^*, dx_1^*] = u \cdot I(C_0^*, C_1^*). \end{aligned} \tag{3.5}$$

Since C_0^* and C_1^* lie in a plane through the origin orthogonal to u, we have that $I(C_0^*, C_1^*)$ is a vector parallel to u, so $u \cdot I(C_0^*, C_1^*) = \|I(C_0^*, C_1^*)\|$, and (3.4) follows.

We remark that the definition of the mixed vector invariant extends in a natural way to piecewise smooth curves, and in particular to polygonal closed curves, and all the properties we have been discussing are satisfied. The results for volumes of nonconvex prismoids that we consider in the next section also hold in this more general framework.

4. Volumes of nonconvex prismoids. A special instance of the preceding section of interest to us is when C_0 and C_1 are simple closed curves lying in parallel planes π_0 and π_1 distance h apart. Assume the parametrizations $x_0(t)$ and $x_1(t)$, $0 \leqq t \leqq T$, are such that the line segments joining $x_0(t)$ to $x_1(t)$ generate a ruled surface which together with the regions in π_0 and π_1, bounded by C_0 and C_1, bound a solid K. Assume also that the parametrizations are consistent with the orientations of π_0 and π_1 so that $I(C_0) = A_0 u$ and $I(C_1) = A_1 u$, where A_0 and A_1 are the areas enclosed by C_0 and C_1, and u is a unit vector orthogonal to π_0 and π_1. Then for fixed z, $0 \leqq z \leqq h$, the plane curve C given by $x(t) = (1 - \frac{z}{h})x_0(t) + (\frac{z}{h})x_1(t)$, bounds the cross-section of the solid K by a parallel plane between π_0 and π_1 at distance z from π_0. If $F(z)$ is the area of this cross-section, we have $I(C) = F(z)u$, and by (3.3)

$$\begin{aligned} F(z)u &= \left(1 - \frac{z}{h}\right)^2 I(C_0) + 2\left(\frac{z}{h}\right)\left(1 - \frac{z}{h}\right) I(C_0, C_1) + \left(\frac{z}{h}\right)^2 I(C_1) = \\ &= \left[\left(1 - \frac{z}{h}\right)^2 A_0 + 2\left(\frac{z}{h}\right)\left(1 - \frac{z}{h}\right) A_{01} + \left(\frac{z}{h}\right)^2 A_1\right] u, \end{aligned} \tag{4.1}$$

where we define A_{01} by the relation $I(C_0, C_1) = A_{01}u$. Now we see that the volume of K is given by integrating $F(z)$ over the interval $[0, h]$, and from (4.1) it is immediate that this volume $V(K)$ is given by exactly the same formula (1.4),

$$V(K) = \frac{h}{3}(A_0 + A_{01} + A_1), \tag{4.2}$$

except in this case the top and bottom faces of K *need not be convex*, and also the solid K *need not be convex*, and A_{01} is a "generalized mixed area" depending on the parametrizations of C_0 and C_1.

Furthermore, since in (4.1) $F(z)$ is a quadratic function of z, we see that Simpson's rule for approximation of integrals gives the exact result and we may conclude immediately that the prismoidal formula

$$V(K) = \frac{h}{6}(A_0 + 4M + A_1) \tag{4.3}$$

also holds in this more general case, where M is still the area of the cross-section of K by a parallel plane midway between π_0 and π_1. Of course this result could also be alternatively derived as in the case of convex curves, using (4.2) and the bilinearity of $I(C_0, C_1)$.

To better understand the relationship of the general case to the special case of convex curves, note that when C_0 and C_1 are smooth strictly convex curves parametrized in such a manner that corresponding outward normal vectors (in π_0 and π_1, respectively) point in the same direction, then the rulings of the lateral surface of the solid K are the intersections of K with lateral supporting planes. This corresponds to the fact that if x_0 and x_1 are boundary points of convex sets K_0 and K_1, having outward normal vectors pointing in the same direction, then $(1-\alpha)x_0 + \alpha x_1$ is a boundary point of $(1-\alpha)K_0 + \alpha K_1$, $0 \leqq \alpha \leqq 1$, with outward normal in the same direction.

5. The twisted truncated pyramid. We illustrate the general volume formulas in the preceding section in a special case of twisted pyramids.

We begin with a cone based on an arbitrary closed curve C_0 in a horizontal plane π_0 and truncate the cone with a parallel plane π_1 at distance h from π_0. The cross-section is a curve similar to C_0. We denote the ratio of similarity by β. We imagine the lateral surface of the truncated cone to be formed by stretching elastic threads from the points of the upper curve to the corresponding points of C_0 along the generators of the cone.

Now rotate the top curve through an angle θ about some vertical axis and denote the resulting curve by C_1. Then the ruled lateral surface formed by the twisted threads, together with the top and bottom bases, determine a nonconvex prismoid, a twisted truncated cone.

It can be shown that if this solid is cut by a horizontal plane whose height is in the ratio α with h, then the cross-section is similar to C_0, with similarity ratio $[\alpha^2 + \beta^2(1-\alpha)^2 + 2\beta\alpha(1-\alpha)\cos\theta]^{1/2}$ and rotated through angle $\arctan[\alpha \sin\theta / (\beta(1-\alpha) + \alpha\cos\theta)]$. The mid-section has the area $M = \frac{1}{4}(1 + 2\beta\cos\theta + \beta^2)A_0$. The generalized mixed area $A_{01} = A(C_0, C_1)$ can be shown to be

$$A_{01} = \beta A_0 \cos\theta, \tag{5.1}$$

where A_0 is the area enclosed by C_0. Here one keeps in mind that the parametrizations of C_0 and C_1 are the natural ones induced by the translation, size change, and rotation through θ that takes C_0 onto C_1, the parametrizations in which corresponding points are connected by the twisted elastic threads.

We remark that in case the bases of a nonconvex prismoid (as described in Section 4) are two arbitrary polygons, each parametrized in a piecewise linear fashion, then the horizontal cross-sections are again polygons (although the ruled surface need not be polyhedral).

From either (4.2) or (4.3) we obtain for the volume V of the twisted truncated cone

$$V = \frac{h}{3} A_0 (1 + \beta \cos\theta + \beta^2). \tag{5.2}$$

In case $\beta = 1$, so the top and bottom bases are congruent, we have

$$V = \frac{hA_0}{3}(2 + \cos\theta) = hA_0 \cos^2\frac{\theta}{2} + \frac{hA_0}{3}\sin^2\frac{\theta}{2}, \tag{5.3}$$

exhibiting the volume as the sum of the volumes of the inscribed cylinder and an outer "tangential" double cone. This illustrates an alternate approach to finding such volumes introduced by Mnatsakanian [5] and applies to the general case of the twisted truncated cone.

Finally consider the case where C_0 is a square of sidelength b and C_1 a square of sidelength a turned through angle θ. We now have a twisted truncated square pyramid and, since $\beta = a/b$, the formula (5.2) gives for the volume

$$V = \frac{h}{3}(a^2 + ab\cos\theta + b^2). \tag{5.4}$$

If the twist angle θ is $180°$, the result is a double cone with square bases, that is, the union of two similar square pyramids whose heights sum to h and whose bases have sidelengths a and b, respectively; this has volume

$$V = \frac{h}{3}(a^2 - ab + b^2).$$

Untwisting corresponds to $\theta = 0$ and brings us back to where we began, the Greatest Egyptian Pyramid in (1.1).

REFERENCES

[1] Blaschke, W., *Vorlesungen über Integralgeometrie*, I, Hamburger math. Einzelschriften, 20, B. G. Teubner, Leipzig, 1935. *Zbl* **12**, 414–415; 2. erw. Aufl., B. G. Teubner, Leipzig, 1936. *Zbl* **14**, 325–326. *Vorlesungen über Integralgeometrie*, II. Hamburger math. Einzelschriften, 22, B. G. Teubner, Leipzig und Berlin, 1937. *Zbl* **16**, 277–278. The two volumes reprinted in one by Chelsea, New York, N.Y., 1949. 3. Aufl., Deutscher Verlag der Wissenschaften, Berlin, 1955. *MR* **17**-888. *Zbl* **66**, 407–408.

[2] Chakerian, G. D. and Goodey, P. R., Inequalities involving convex sets and their chords, *Convexity and graph theory* (Jerusalem, 1981), North-Holland Math. Studies, **87**; *Annals of Discrete Mathematics* **20** (1984), North-Holland, Amsterdam–New York, 1984, 93–101. *MR* **86k**:52011

[3] Flanders, H., The Steiner point of a closed hypersurface, *Mathematika* **13** (1966), 181–188. *MR* **34** #6695

[4] Heath, T. L., *A history of Greek mathematics*, Vol. II, Clarendon Press, Oxford, 1921. *Jb. Fortschritte Math.* **48**, 46

[5] Mnatsakanian, M. A., On the area of the region on a developable surface, *Akad. Nauk Armyan. SSR Dokl.* **73** (1981), 97–102 (in Russian). *MR* **83d**:53006

[6] Polya, G., *Mathematical discovery*, John Wiley & Sons, New York, 1981. *MR* **82f**: 00002

[7] Waerden, B.L. van der, *Science awakening*, P. Noordhoff Ltd., Groningen, 1954. *MR* **16**–1; 2nd ed., Oxford Univ. Press, New York, 1961. *MR* **23** #A2291; John Wiley and Sons, New York, 1963.

(Received January 10, 1994)

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
DAVIS, CA 95616
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 31 (1996), 455–478

ISOPERIMETRIC NETWORKS IN THE EUCLIDEAN PLANE

M. N. BLEICHER

Dedicated to László Fejes Tóth on his eightieth birthday

Abstract

In this paper we consider the following problem. Given n positive numbers, $(a_1, a_2, \ldots, a_n)$ find the network of least total arc length such that the bounded components of its complement can be put into $n+1$ regions, R_i, $0 \leqq i \leqq n$, such that the sum of the areas of the components in region R_i is a_i for $i > 0$. The components in R_0 are called dead spaces.

It is widely conjectured that in the length minimizing network, there are no dead spaces and the regions are connected. The conjecture is known to be true only for the case $n=2$. On the assumption that the regions are connected, it is known only for $n \leqq 3$ that there are no dead spaces. In this paper, using novel methods, we prove the known that there are no dead spaces for 3 connected regions. Next we use those methods to prove that for 4 connected regions there are no dead spaces. It is hoped that these new methods will enable us to attack the problems of more regions and not necessarily connected regions.

1. Introduction

The problem of dividing the plane or space into a finite number of regions with prescribed area or volume using the least amount of material, arc length or surface area, arose in antiquity. It comes from both theoretical and practical problems, including the study of soap bubbles and their two-dimensional analogue, see [3, 4, 10, 11, 12, 19, 20, 22].

The solution of the problem of one region both unrestricted and restricted by variety of side conditions was known in antiquity, as evidenced by Zenodorus's book "περὶ ἰσομέτρων σχημάτων" on isoperimetric problems. The existence of solutions was assumed in antiquity but not rigorously proven, say by means of the Blaschke Selection Theorem, [7], or calculus of variations techniques, [1], until modern times. This problem has been generalized to many contexts. See [3, 12, 17, 18, 19] for more information and further references.

In this paper we consider the problem in two-dimensional Euclidean space. While it seems intuitively clear that for an optimal configuration the regions will be connected and there will be no dead spaces which while surrounded are not part of any region, this has never been proven. It is

---

1991 *Mathematics Subject Classification.* Primary 52A38; Secondary 52A40.
Key words and phrases. Isoperimetry, soap bubbles, isoperimetric inequalities.

known only if the number of regions is two. These notions will be defined precisely in the next section. Many authors in the very definition of the problem incorporated those assumptions, which are often required as additional side conditions in applications, [3, 10, 11, 12]. More recently the problem has been discussed in full generality, [5, 15, 18].

In Section 2, we prove some general results on the nature of the edge network, the dead spaces, the unbounded component, and the relationships between components.

In Section 3, using the results of Section 2, we show that for connected regions, if the number of regions is at most four, then there are no dead spaces. For four regions this result is new; for three regions this method gives a different proof.

In Section 4, we give some conjectures and unsolved problems.

References may be found in Section 5.

2. Fundamentals

DEFINITION 1. A *division network* or, briefly, a *network* is a finite union of rectifiable arcs which divide the plane into a finite number of components, more precisely the components of the complement of the network in the plane, exactly one of which is unbounded.

DEFINITION 2. The *length* of the network is the sum of the lengths of all the arcs of which it is composed.

DEFINITION 3. A *region* is a finite union of disjoint bounded components.

DEFINITION 4. A network *satisfies the parameters* $(a_1, a_2, \dots, a_n)$ iff the bounded components can be partitioned into $n+1$ regions R_j, $0 \leqq j \leqq n$ such that for each j, $0 < j \leqq n$ the sum of the areas of the components assigned to R_j, is a_j. The regions assigned to R_0 are called *dead spaces* [15, 5, 18]; the unbounded component together with those in R_0 are called *free spaces* in that we are free to change the area without changing the fact that the network satisfies the given parameters. The bounded components assigned to regions other than R_0 are called *encumbered* components.

The following notational convention will be used in this paper:

NOTATIONS. 1. D_0 will denote the (unique) infinite component.

2. D_i, $0 < i \leqq k$, will denote the dead spaces, where k is the number of components assigned to R_0.

3. Encumbered components be denoted by E_j or E_j^m to emphasize it is a bounded component of some region R_m with $0 < m$. Thus the components are in two disjoint classes, free and encumbered.

4. B_i will denote a bounded component which may or may not be free.

5. C_i will denote an arbitrary component.

6. Two components or regions are called *separated* from one another iff their closures have no common points.

REMARKS. 1. The area of any or all of the components of the free region is unimportant in determining whether or not a network satisfies a given set of parameters.

2. The same network could satisfy more than one collection of parameters with a different assignment of components to regions.

DEFINITION 5. A network is *optimal for the parameters* $(a_1, a_2, \ldots, a_k)$ or simply *optimal* iff there is no shorter network satisfying the given parameters.

Given an optimal network for some parameters, if there are q bounded components B_i each of area b_i, $0 < i \leq q$, then the network is also optimal for the parameters $(b_1, b_2, \ldots, b_q)$. But as a network satisfying $(b_1, b_2, \ldots, b_q)$, it is an optimal network with connected regions and no dead spaces. Such a network is of the type studied in [3] and all the theorems of that paper apply; some of these results can also be found in [15]. Restating these previous results, we obtain the following theorem.

THEOREM 1. *A network that does not satisfy all the following optimality conditions can be replaced with a shorter network with the same area divisions of its bounded components. Optimality conditions:*

1. *All the arcs are segments of circles or lines; i.e., they have constant curvature.*

2. *All nodes in the graph of arcs meet exactly three edges, i.e., all nodes have degree three.*

3. *The sum of the signed curvature of the three edges at any node is zero.*

4. *The angle between any two arcs at the same node is* $\frac{2\pi}{3}$.

5. *Let* q *be the number of bounded components. The number of nodes is* $2q - 2$ *and the number of edges* $3q - 3$.

6. *The network is connected.*

LEMMA 1. *Given an optimal network and a simple closed curve,* Γ, *meeting the network finitely often and containing no nodes of the network, the signed sum of the curvatures of all the edges crossed by* Γ, *counting multiplicities, is zero.*

PROOF. This can be proven easily, by induction on the number of vertices surrounded by Γ, using Part 3 of Theorem 1. An alternate proof can be found in [5]. □

The next lemma, a consequence of Euler's Formula, is a useful, known result for finite trihedral graphs [6, 16] or for a straightforward proof of a special case of the dual, see [2, p. 21].

Lemma 2. *If f_i denotes the number of components with i edges, then*

$$\sum_{i=1}^{N}(6-i)f_i = 12,$$

where N is the maximum number of edges of any component.

We shall see shortly that for optimal networks with $n > 2$, $f_i = 0$ for $i < 3$.

Lemma 3. *In an optimal network free spaces are separated. I.e., for $0 \leqq i < j \leqq k$ we have $\overline{D}_i \cap \overline{D}_j = \emptyset$.*

Proof. Since all nodes are of degree three, two components with a common node have a common edge. Thus either the two free spaces are separated or they have at least one common edge. If they are not separated, erasing this common edge yields a new shorter network satisfying the same parameters, which contradicts the optimality. □

Lemma 4. *In an optimal network two distinct components of the same region are separated.*

Proof. As in the previous proof we erase the common edge and obtain a better network for the same parameters, contradicting the optimality. □

Lemma 5. *In an optimal network, if two regions, necessarily distinct, meet in two distinct edges then the two distinct edges have the same curvature when both are viewed from the same region.*

Proof. Let e and e' be the two edges at which R and R' meet. Let e and e' have curvature κ and κ' when viewed from region R. Suppose $\kappa \neq \kappa'$.

First suppose that the curvatures have opposite signs. On each arc draw small chords which cut off equal areas. Replace the arcs by these chords. The area of each region is unchanged, but the network length is shortened, violating the optimality.

We now suppose that κ and κ' do not have opposite signs, but one might be zero. From each edge take a small length of arc with endpoints equidistant on each piece and interchange them to obtain new edges e_1 and e'_1. Neither the area of the regions nor the total arc length of the network are changed by this exchange, but neither of the new edges have constant curvature, and thus can be replaced with shorter edges yielding the same area division by Theorem 1, again violating optimality. □

Lemma 6. *In an optimal network two distinct components cannot have two edges in common.*

Proof. We may suppose we have two components E and C since they cannot be from the same region. Let e and e' be the common edges. Let the endpoints of e and e' be A and B and A' and B', respectively, chosen so that the order walking around the boundary of C with positive orientation is $ABB'A'$ (see Figure 1). Neither $A = A'$ nor $B = B'$ are possible since the graph is trivalent and the edges are distinct. We now make the argument for the case when C is a bounded component, it is actually simpler when C is unbounded.

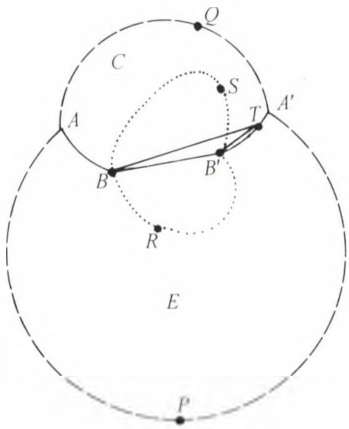

Fig. 1

Choose P on the part of bounding curve of E which joins the A and A', but does not contain the points B or B'. We denote this curve by APA'. Similarly choose Q so that AQA is the part of the boundary of C joining the points A and A', but does not contain B or B'. Choose R and S on the part of the boundary curves of E and C, respectively, which join the points B and B', but do not contain A or A'. Thus $ABRB'A'PA$ is the boundary of E, while $ABSB'A'QA$ is the boundary of C (in the opposite orientation).

We now describe how the region surrounded by $BRB'SB$ may be moved "continuously" from A toward A'. Let δ be chosen as a positive real number not exceeding the length of the arc $B'A'$. Choose T on $B'A'$ so that the arc length of $B'T$ is δ. Consider the triangle, $\Delta BB'T$, possibly degenerate. Consider the mirror image of the triangle in the perpendicular bisector of the base $\overline{BT}$ of this triangle. To the mirror image of the line segment $\overline{BB'}$ attach all the edges, not the mirror image, which were connected to BB' without using either of the edges AB or $A'B'$ to make the connection. (This set of edges is indicated by dotted lines in Figure 1.) This exchange effects neither the total length of the edges nor the area divisions. Unfortunately, the new graph may have edges which overlap other parts of the graph. However, for small enough δ this cannot happen since there is a minimal positive distance between the closed arcs to be exchanged and the unmoved arcs APA' and AQA' and as $\delta \to 0$ the exchanged arcs tend to their original position. Two cases need to be distinguished, the first when the graphs do not overlap for $T = A'$ and the second when they do overlap for $T = A'$.

In the first case the new configuration has a vertex of degree 4 at A' and hence by Theorem 1 can be improved; this contradicts the optimality of the given network. In the second case there will be a value of δ for which the exchanged arcs are tangent to the stationary arcs, which, at the point of tangency, leads to a node of degree greater than three. Again this violates the optimality. □

LEMMA 7. *In an optimal network any path joining two of the free components cannot cross arcs of the network all of which have positive curvature.*

PROOF. Suppose such a path exists. There is a positive number, ε, such that it is possible, on each arc the path crosses, to cut off a segment of the circular arc with area ε. Replacing the arcs of these segments by the chords of these segments in the network, creates a new network of shorter length, but has left unchanged the area of any of the components except the free ones at each end of the path. But changing the area of only free components leaves the same set of parameters satisfied, but with shorter total length, which violates the optimality. □

REMARK. Of course there can be no such path which crosses only arcs of negative curvature, since traversing the path in the opposite direction it crosses only positive curvature.

The next lemma is one we shall use repeatedly in the analysis of permissible network in both this and the next section.

LEMMA 8. *At any node at which no edge has zero curvature, one of the three components meeting there has both edges of positive curvature, one has both negative, and the third has one positive and one negative.*

PROOF. Since each edge is viewed from both sides, the number of positive and negative curvatures are equal. It follows that either each region has one positive and one negative edge at the node or the lemma is true.

If they each have one positive and one negative, then the edges must all curve the same direction on leaving the node. But that is impossible by Theorem 1 since the sum of the curvatures would not be zero. The lemma follows. □

DEFINITION 6. If C is any component of a network we define $e(C)$ to be the *number of edges* of that component.

LEMMA 9. *For an optimal network:*

1. *If τ_i is the total curvature of the i^{th} edge of a bounded component, B, oriented from inside that component, then*

$$\sum_{i=1}^{e(B)} \tau_i = 2\pi - \frac{e(B)\pi}{3}.$$

2. *If τ_i is the total curvature of the i^{th} edge of the unbounded component, oriented from the inside of the complement of the unbounded component,*

then

$$\sum_{i=1}^{e(D_0)} \tau_i = 2\pi + \frac{e(D_0)\pi}{3}.$$

3. $e(D_0) = 1$ *iff there is a component, C, for which $e(C) = 1$ iff $n = 1$.*
4. $e(D_0) = 2$ *iff there is a component, C, for which $e(C) = 2$ iff $n = 2$.*
5. $e(D_0) \geqq 3$ *iff for every component, C, $e(C) \geqq 3$ iff $n \geqq 3$.*

PROOF. 1. Since the boundary is a simple closed curve the total curvature around the component is 2π. The exterior angle where any two edges meet is $\pi/3$. This leads to the equation

$$\sum_{i=1}^{e(B)} (\tau_i + \pi/3) = 2\pi,$$

which yields the result.

2. For the unbounded component the exterior angles, when viewed from inside the complement of the unbounded component, are all $-\pi/3$, and the rest is the same as the last proof.

3. Assuming $n = 1$, this is just the classic Isoperimetric Theorem that a circle is the most efficient way to enclose an area. Conversely if $e(D_0) = 1$ this edge must be a circular arc, but since it cannot meet itself at an angle of $2\pi/3$, it must be a complete circle. Since the network must be connected by Theorem 1, we conclude that the circle is the entire network, which thus has only one bounded component.

4. If $e(D_0) = 2$ then clearly there is an E with $e(E) = 2$. Next, suppose for some component, E, $e(E) = 2$, then the boundary of E consists of two circular arcs meeting at, say, P and Q. Since each vertex is trivalent, at each of P and Q there is an edge separating the same two components. By Lemma 6 these edges at P and Q must be the same edge. Since the network is connected, the entire network must consist of these three edges which surround two bounded components, which by Lemma 4 must belong to different regions; i.e., $n = 2$.

Finally, suppose that $n = 2$. We wish to show that $e(D_0) = 2$. This would follow from the fact that the two regions are connected. Proving this statement, which is Main Theorem of [15], is the hard part of the proof, so we refer the reader to the cited paper where this implication is proven.

5. This is just the contrapositive of the last two statements. □

LEMMA 10. *If $e(D_0) = 3$, then all the edges of D_0 have positive curvature when viewed from inside the network.*

PROOF. From Lemma 9, Part 2, we see that at least two of the τ_i must be positive, since their sum is 3π and no one of them can exceed 2π. But τ_i and κ_i have the same sign, so two of the curvatures are positive. Suppose the third edge, say e_3, does not have positive curvature, while e_1 and e_2 do. If the edge e_1 or e_2 is extended on the same circular arcs until they meet again, giving arcs a_1 and a_2 of total curvature α_1 and α_2 which must sum to $8\pi/3$ from Lemma 9, Part 2. But $\tau_1 + \tau_2 = 3\pi + |\tau_3|$ and the arcs e_1 and e_2 only intersect once, which tells us that one of the edges, say e_1, is shorter than the arc a_1 while the other edge, e_2 is longer than the arc a_2. Suppose now that $\tau_3 = 0$. Let $\tau_1 = \alpha_1 - \varepsilon$, then $\tau_2 = \alpha_2 + \varepsilon + \pi/3 = 3\pi - (\alpha_1 - \varepsilon)$. At the end of e_1 draw the inner normal to edge e_3. At the end of e_2 draw the normal to the edge e_3. The lines perpendicular to these normals at these endpoints are distinct, and hence there can be no edge e_3 of zero curvature. Suppose now that $\tau_3 < 0$, then $\tau_2 = 3\pi + |\tau_3| - (\alpha - \varepsilon)$. Thus a_2 is extended by exactly the amount of the total curvature of the edge e_3, thus the inner ray to the circular arc e_3 is parallel to the inner normal to the circular arc e_3 at its other end, but in opposite directions. Thus the edge e_3 is a semicircle. Hence $\tau_1 + \tau_2 = 3\pi + \tau_3 = 4\pi$. But this is impossible since both τ_1 and τ_2 are each less than 2π. Thus it is not possible for the third edge to have negative curvature. □

LEMMA 11. *If $e(D_0) = 4$ and two adjacent edges of D_0 have positive curvature then at least three of the edges of D_0 have positive curvature when viewed from inside the network.*

PROOF. This proof is similar to, but slightly more complicated than, the proof of Lemma 10, since one must look at the normals to the ends of the edge e_4, which will be parallel. □

3. Connected regions

In this section we prove some additional lemmata which hold for optimal networks in which the encumbered regions are required to be connected, but in which dead spaces are allowed. We then consider the cases of 3 or 4 connected regions and show that there are no dead spaces. For 2 or 3 regions this is already known; see [15] or [5, 18], respectively. The proof we give for $n = 3$ is different and is a good warm-up for the proof of the previously unsolved case of $n = 4$ connected regions.

LEMMA 12. *If D is a free region in a graph with n encumbered regions, all of which are connected, then $e(D) \leq n$.*

PROOF. By Lemma 3 each edges of D meets one of the n connected regions. By Lemma 6 they are all from different connected regions. The lemma follows. □

NOTATIONS. 1. The number of free components with i edges is denoted by $\tilde{f}_i$.

2. The number of encumbered regions with i edges is denoted by $\hat{f}_i$. Thus, $f_i = \tilde{f}_i + \hat{f}_i$.

3. An edge separating a free region from another region is called a *dead* edge. An edge separating two encumbered regions is called a *live* edge.

4. The number of dead edges of the component, E or E_i, is denoted by d or d_i, respectively.

Since the free components are separated, every edge is either a dead edge or a live edge.

LEMMA 13. *If E has e edges then $d \leqq \left[\frac{e}{2}\right]$.*

PROOF. Since free regions are separated, between every two dead edges there must be a live edge, but live edges need not be separated, therefore at least half the edges are live. Hence the number of dead edges is at most the greatest integer in half of number of edges of the component. □

LEMMA 14. *In an optimal network of n regions, all of which are connected, the number of edges of a bounded region is at most $2(n-1)$; i.e., $e(E) \leqq 2(n-1)$.*

PROOF. Since two components can have only one common edge, the number of live edges of a bounded region is at most $(n-1)$. Since the dead edges of E must be separated by one or more live edges, there can be at most $(n-1)$ of them. Since every edge is either live or dead, the inequality follows. □

THEOREM 2. *An optimal network for three connected regions has no dead spaces; in fact, it must have the topological form and curvatures of the network in Figure 3.*

PROOF. By Lemma 9, Part 5 all components have at least three edges. By Lemma 12 the free components have at most 3 edges; hence all dead spaces and the unbounded component have three edges. On the other hand by Lemma 14 no bounded region, which by hypothesis is one component, can have more than 4 edges. It follows that $f_i = 0$ for $i \neq 3, 4$. Lemma 2 yields,

$$3f_3 + 2f_4 = 12. \tag{1}$$

Since all the free spaces have three edges and there is always at least one free space, the unbounded component, we see that $f_3 \geqq 1$. With that restriction the only non-negative integral solutions to Equation 1 are:

$$(f_3, f_4) = (2, 3) \quad \text{or} \quad (4, 0).$$

We consider the solution $f_3 = 2$, $f_4 = 3$ first. Suppose there is such an optimal network. Since there are exactly three encumbered components, one for each region, there must be two free components, D_0 and D_1 which by

Lemma 12 have three edges each. It follows that the encumbered components each have four edges. By Lemma 10 all the edges of D_0 have positive curvature viewed from inside the network. But all since the dead edges of each E_i have the same curvature, all the edges of D_1 have negative curvature viewed from D_1. But by Lemma 9, Part 1, the curvatures cannot all be negative, a contradiction. This solution cannot lead to an optimal network.

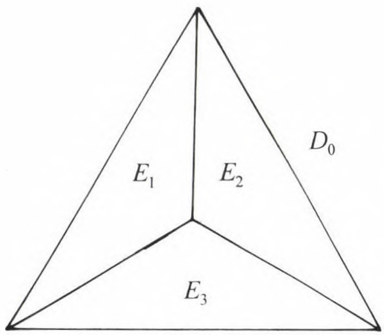

Fig. 2

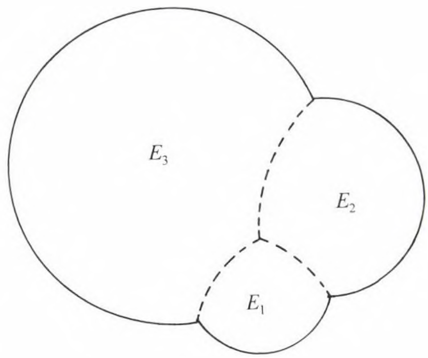

Fig. 3

Consider the other solution, $f_3 = 4$, $f_4 = 0$ and these are the only faces. Three of these four components are the three encumbered regions; the fourth is the unbounded component. The graph of edges is the Schlegel diagram of the tetrahedron, Figure 2. By Lemma 10 all the edges of D_0 have positive curvature. If κ_i is the curvature of the dead edge of E_i numbered so that $\kappa_1 \geqq \kappa_2 \geqq \kappa_3 > 0$ then Figure 3 is the only possible configuration, with the proviso that the edge separating E_i from E_j has zero curvature precisely when $\kappa_i = \kappa_j$, otherwise the sign of the curvature is as pictured. □

THEOREM 3. *An optimal network for four connected regions has no dead spaces; in fact it must have the topological form and the edge curvature pattern of one of the networks in Figure 4.*

PROOF. Let D be a free component, from Lemma 9 and Lemma 12 we conclude $3 \leqq e(D) \leqq 4$. For an encumbered component E we conclude $3 \leqq e(E) \leqq 6$. Lemma 2 yields the equation

$$3f_3 + 2f_4 + f_5 = 12, \tag{2}$$

where we also know

$$f_3 + f_4 \geqq k + 1 \quad \text{and} \quad f_3 + f_4 + f_5 + f_6 = f = k + 5, \tag{3}$$

and thus

$$f_5 + f_6 \leqq 4, \quad f_5 \leqq 4, \quad \text{and} \quad f_3 + f_4 \geqq 1. \tag{4}$$

Furthermore,

$$\widehat{f}_3+\widehat{f}_4+\widehat{f}_5+\widehat{f}_6=4, \quad f_5=\widehat{f}_5, \quad f_6=\widehat{f}_6. \tag{5}$$

The complete set of non-negative integral solutions to equation (2) satisfying the inequalities (4) are:

$$\begin{aligned}(f_3,f_4,f_5) &=(4,0,0)\\ &=(3,1,1) \quad \text{or} \quad (3,0,3)\\ &=(2,3,0) \quad \text{or} \quad (2,2,2) \quad \text{or} \quad (2,1,4)\\ &=(1,4,1) \quad \text{or} \quad (1,3,3)\\ &=(0,6,0) \quad \text{or} \quad (0,5,2) \quad \text{or} \quad (0,4,4).\end{aligned}$$

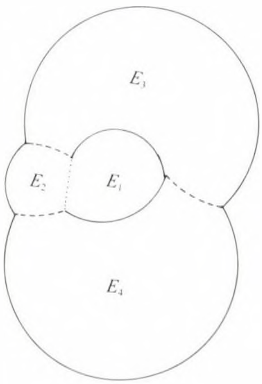

Fig. 4a (i)

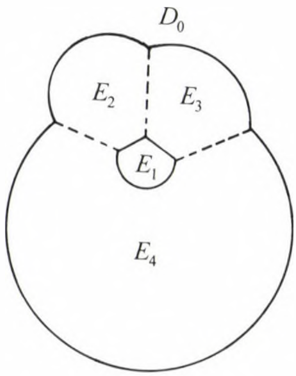

Fig. 4a (ii)

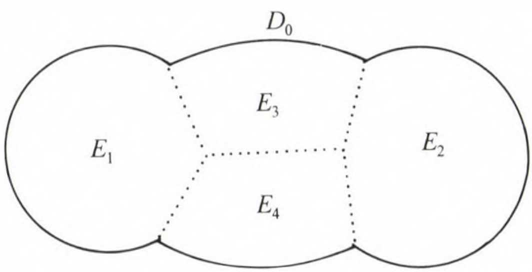

Fig. 4b

Before analyzing these 11 cases, we note two Lemmata.

LEMMA 15. *If* $f_6>0$, *i.e., there is a component* C, *with* $e(C)=6$, *then:*
1. C *is encumbered.*
2. C *meets all other encumbered components and three distinct free components.*

3. *Every other encumbered component has at least two dead edges and at least four edges.*

4. $k \geqq 2$ *and* $\hat{f}_3 = 0$.

Proof. Since free components have at most 4 edges by Lemma 12, we deduce that C is encumbered. Since there are only three other encumbered components, there must be at least three dead edges which alternate (Lemma 3) and which are from distinct components (Lemma 6). Since every other encumbered component meets E between two free components and these free components cannot meet, there must be at least one more edge on the other encumbered component. The inequalities follow, which completes the proof of the lemma. □

Lemma 16. *If $f_5 > 0$, i.e., there is a component C with $e(C) = 5$ then C is an encumbered component, $k > 0$ and $f_6 \leqq 1$.*

Proof. By Lemma 12, C is encumbered. Since there are only three other encumbered components, two of the edges of C must meet, necessarily distinct, free components; so $k > 0$. At one vertex of C two other encumbered components must meet. Then each of these other components have two adjacent edges which meet encumbered components and hence neither of these components can have 6 edges. Since these two components, together with C account for three of the four encumbered components, there can be at most one component with six edge, which implies the lemma. □

We return to the proof of the theorem.

We now consider the cases.

Case 1. $(f_3, f_4, f_5) = (4, 0, 0)$.

By equation (3) we know that $f_6 > 0$. From Lemma 15 we see that $k \geqq 2$ and from the first item of inequalities (3) that $k \leqq 3$. Thus there are two possibilities: $k = 2$ and $f_6 = 3$; or $k = 3$ and $f_6 = 4$.

If $k = 2$, then there are three free components each with three edges and four encumbered component, three with six edges and one with three edges. This violates Lemma 15, Part 4.

If $k = 3$ the configuration is combinatorially possible; it is the Schlegel diagram of the snub tetrahedron, Figure 5. However, as we show next, when curvature conditions are considered, it cannot come from an optimal network.

In Figure 6, we follow the implications of our knowledge of curvature of the edges of an optimal graph. Since all the free spaces have the same number of edge, we draw the graph so the outside face is D_0. All the edges of D_0 must have positive curvature when viewed from inside the graph (Lemma 10). Lines drawn solidly are known to have the indicated non-zero curvature, line drawn with dashes have the indicated curvature except that it might be zero, and dotted lines have unknown curvature. This enables us to draw the solid lines labeled 1. Since all edges between E_i and any free component have the same curvature when viewed from E_i we can draw the solid lines marked 2. By Lemma 9, Part 1, all the components have an edge with

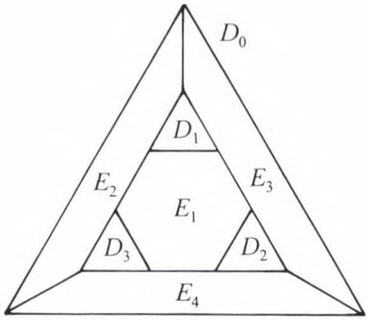

Fig. 5

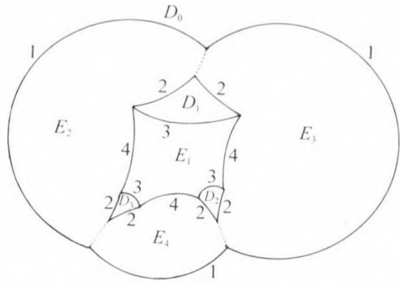

Fig. 6

positive curvature viewed from inside the component. This enables us to draw the solid line marked 3. From Theorem 1, Part 3, we see the curvature of the edges marked 4. Thus E_4 has only edges with negative curvature, a contradiction to Lemma 9, Part 1. Thus the configuration of Figure 5 cannot occur in an optimal network.

Case 2. $(f_3, f_4, f_5) = (3, 1, 1)$.

Since $f_5 > 0$, we see fromLemma 16 that $k > 0$. It follows from equation (3) that $f_6 > 0$, from which, via Lemma 15, we see that both $k \geqq 2$ and $\widehat{f}_3 = 0$. Again by equation (3), we see that $f_6 = k$ and from inequality (4) that $f_6 \leqq 3$. For $k = f_6 = 2$, we see that $\widetilde{f}_3 = 3$, $\widetilde{f}_4 = 0$, $\widehat{f}_3 = 0$, $\widehat{f}_4 = f_4 = 1$, $\widehat{f}_5 = 1$, and $\widehat{f}_6 = 2$. Since the component E with four edges meets the component with six edges, it must have two dead edges. Using Lemma 15, Part 3 to count the total number of dead edges two ways, we obtain

$$9 = \sum_{0}^{2} e(D_i) = 2f_4 + 2f_5 + 3f_6 = 10,$$

a contradiction. For $k = f_6 = 3$ we obtain, in a similar way,

$$13 = \sum_{0}^{3} e(D_i) = 2f_4 + 2f_5 + 3f_6 = 0 + 2 + 9 = 11.$$

Hence Case 2 is impossible.

Case 3. $(f_3, f_4, f_5) = (3, 0, 3)$.

Again, we see from Lemma 16 that $k > 0$. Also from inequalities (4) we deduce $f_6 \leqq 1$ and $f_6 = k - 1$. Thus the only possibilities are: $k = 1$ and $f_6 = 0$; or $k = 2$ and $f_6 = 1$. If $k = 1$, there are exactly six dead edges, three from each free region. Since each pentagonal encumbered component has two dead edges, the three edged encumbered component, say E_4 has none.

Thus E_4 meets each E_i for $1 \leqq i \leqq 3$. Hence each E_i, $i > 0$ meets every other encumbered component on three adjacent edges. This means the two dead edges are adjacent in the five-edged encumbered components, contradicting Lemma 3.

Combinatorially $k = 2$ and $f_6 = 1$ is possible, see Figure 7, the Schlegel diagram of a tetrahedron with three corners clipped. We now use curvature arguments to show it is impossible for an optimal network. We know that $e(D_0) = 3$ as pictured. From Lemma 10 we see all the edges of D_0 have positive curvature, Figure 8. Deducing the curvature of the edges in the order indicated by the labels, we again arrive at a contradiction of a face with all negative curvature edges. This case does not occur in an optimal network.

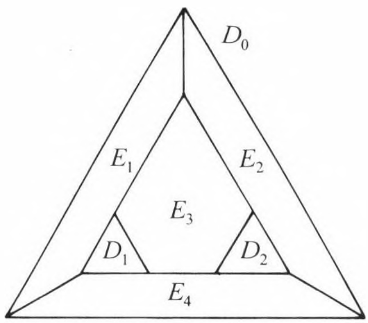

Fig. 7

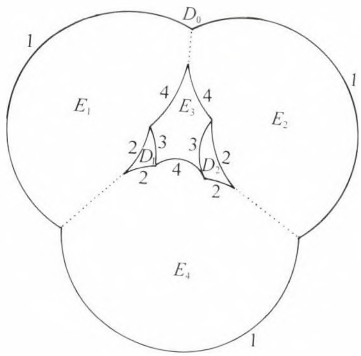

Fig. 8

Case 4. $(f_3, f_4, f_5) = (2, 3, 0)$.

If $k = 0$, then D_0 is the only free component. Depending upon whether $e(D_0) = 3$ or 4, we get the two possible configurations which may actually occur as solutions, Figure 9, which are both Schlegel diagrams for the triangular prism. At the end of the proof we show that the signs of the curvature of the edges are as indicated in Figure 4.

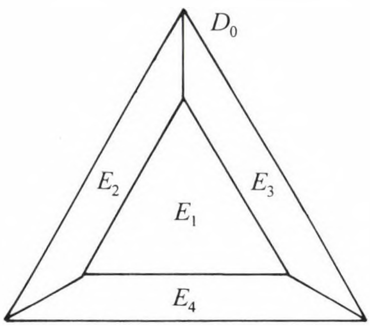

Fig. 9a

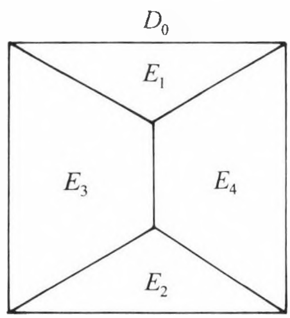

Fig. 9b

If $k > 0$, then $f_6 = k > 0$ from equation (3). Thus Lemma 15 implies

$k \geq 2$, $\hat{f}_3 = 0$ and each of the encumbered components with four edges has exactly two dead edges. We now count dead edges from two perspectives

$$\sum e(D_i) = 2 \cdot 3 + (k-1)4 = 2\hat{f}_5 + 3\hat{f}_6 = 2(4-k) + 3k.$$

Solving for k yields $k = 2$. This is combinatorially realizable, Figure 10, in two ways, both of which are Schlegel diagrams of a triangular prism with two clipped corners at the ends of an edge separating two rectangular faces. However, neither of these graphs are possible optimal networks because of the curvature requirements.

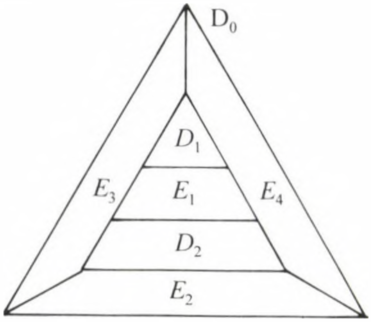

Fig. 10a

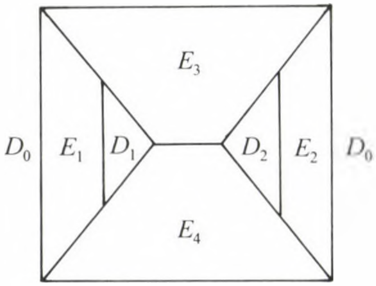

Fig. 10b

If $e(D_0) = 3$, we know all of the edges of D_0 have positive curvature, Figure 11a. We infer that the edges marked 2 have the indicated curvature from Lemma 5. Since the dead spaces must have an edge of positive curvature, we deduce the sign of the curvature of the edges labeled 3. By examining the internal vertices, we conclude the curvature of edges 4 are as indicated. This leads to the contradiction that E_1 has no edges of positive curvature.

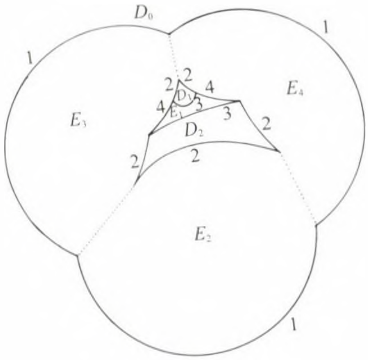

Fig. 11a

If $e(D_0) = 4$, we know at least two of the edges have positive curvature. If exactly two edges of D_0 have positive curvature, then they must be opposite, by Lemma 11. If exactly three edges of positive curvature then again we have an edge of D_0 with non-positive curvature surrounded by edges of D_0 which have positive curvature. This leads to Figures 11b and 11c. In the first E_1 has no edges of positive curvature; in the second, or E_4 will fail to have an edge with positive curvature. These contradictions show it is impossible to have any edges of D_0 with non-positive curvature. If all of the edges of D_0 have positive curvature, the internal dead space will have only edges of negative curvature, a contradiction. Thus, this case cannot arise in an optimal network.

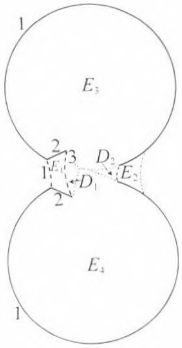

Fig. 11b

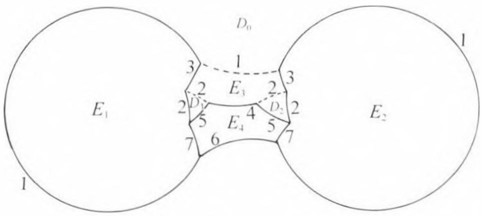

Fig. 11c

Case 5. $(f_3, f_4, f_5) = (2, 2, 2)$.

We first deduce that $f_6 = k - 1$ and $f_6 \leq 2$. Since a free component can have at most four edges, there are at most four of them and $k \leq 3$. Thus three subcases arise: $k = 1$ and $f_6 = 0$; $k = 2$ and $f_6 = 1$; or $k = 3$ and $f_6 = 2$.

Subcase 1. $k = 1$ and $f_6 = 0$.

This network is the Schlegel diagram of a triangular prism with one corner clipped. We note that $\tilde{f}_4 = 2 - \tilde{f}_3$, $\hat{f}_3 = 2 - \tilde{f}_3$, $\hat{f}_4 = 2 - \tilde{f}_4 = \tilde{f}_3$, and $\hat{f}_5 = f_5 = 2$.

Counting dead edges we get

$$3\tilde{f}_3 + 4(2 - \tilde{f}_3) = \sum e(D_i) \leq 1(2 - \tilde{f}_3) + 2\tilde{f}_3 + 2 \cdot 2,$$

with equality iff the three-edged encumbered components have one dead edge and the four-edged ones have 2 dead edges. The inequality implies $1 \leq \tilde{f}_3 \leq 2$ with equality on the left-hand inequality iff the same conditions for the last equality hold.

Thus there are three possibilities for how this network looks. There can be two triangular free spaces, $\tilde{f}_3 = 2$, Figure 12a. Or there can be one triangular and one quadrangular free space, $\tilde{f}_3 = 1$, leading to Figures 12b and 12c in which D_0 has three or four edges, respectively.

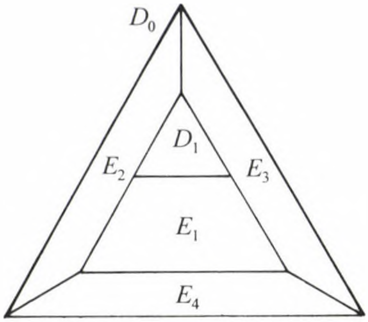

Fig. 12a

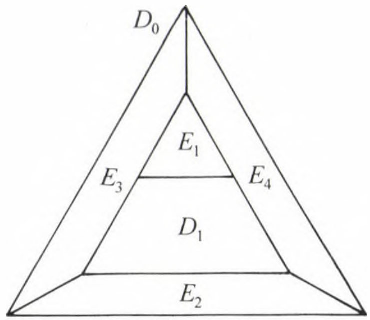

Fig. 12b

Configurations of the type shown in Figure 12a cannot occur. By Lemma 10, all the edges of D_0 have positive curvature. Following the numbered order of the edges we infer that the curvatures must be as shown. There is then the path Γ crossing only positive curvature edges joining two free components. Lemma 7 says the network is not optimal.

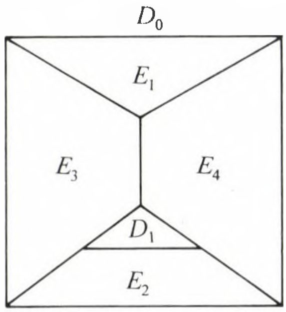

Fig. 12c

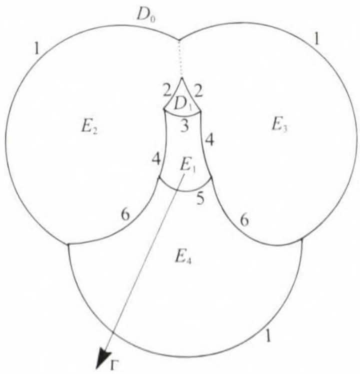

Fig. 13a

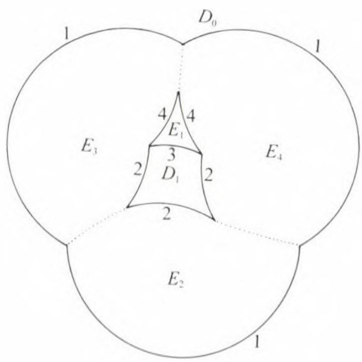

Fig. 13b

In a configuration of the type shown in Figure 12b, we again know that all the edges of D_0 have positive curvature. Figure 13b indicates how to arrive at the contradiction that E_1 has no positive edges. This configuration cannot occur in an optimal network.

From Lemma 10 we infer the configuration of Figure 12c has the following realizations: An edge of D_0 has negative curvature but the adjacent edges of D_0 have positive curvature; or all the edges of D_0 have positive curvature.

In the first realization there are three ways it can be drawn. The surrounded edge of D_0 without positive curvature can be on a component with three, four or five edges. In the first two ways, Figures 13c (i)α and 13c (i)β, we immediately see that the face involved has no edge with positive curvature. The third leads to Figure 13c (i)γ, following the implication for edges as numbered. The deduction for the direction of the edge labeled 4 is that component E_3 must have an edge with positive curvature. The contradiction follows from the fact that all edges of E_4 have negative curvature.

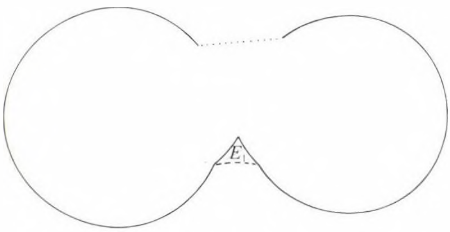

Fig. 13c (i)α

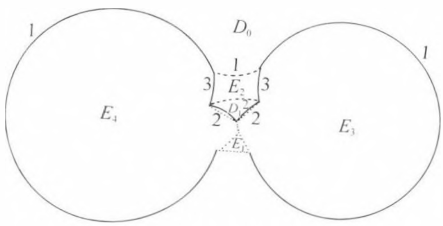

Fig. 13c (i)β

In the second realization all the edges have positive curvature. Since all dead edges of the same component have the same curvature, it follows, Figure 13c (ii) that none of the edges of D_1 have positive curvature, a contradiction. Subcase 1 cannot occur in an optimal network.

Subcase 2. $k = 2$ and $f_6 = 1$.

In this subcase, since $f_6 > 0$, we know, Lemma 15, that for every encumbered component, E, $e(E) \geqq 4$ and that the encumbered component with four edges has two dead edges. Thus counting dead edges two ways we get

$$10 = \sum e(D_i) = 2\widehat{f}_4 + 2f_5 + 3f_6 = 2 + 4 + 3 = 9.$$

This contradiction shows this subcase cannot be optimal.

Subcase 3. $k = 3$ and $f_6 = 2$.

As in the preceding case, we count dead edges.

$$14 = \sum e(D_i) = 2f_5 + 2f_6 = 4 + 6 = 10,$$

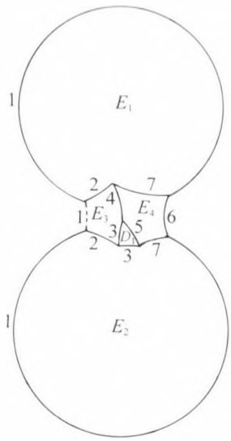

Fig. 13c (i)γ

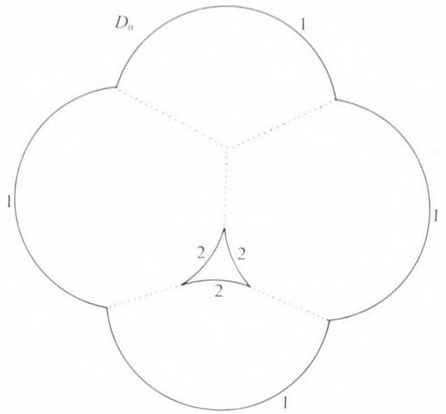

Fig. 13c (ii)

a contradiction, thus this case cannot be optimal.

Case 6. $(f_3, f_4, f_5) = (2, 1, 4)$.

Since free components have at most 4 edges the four components with five edges each are the four encumbered components. Counting dead edges, we obtain the contradiction

$$10 = \sum e(D) = 2f_5 = 8.$$

This case cannot be optimal.

Case 7. $(f_3, f_4, f_5) = (1, 4, 1)$.

We first note from equation (3) that $f_6 + 1 = k$. Also $f_6 \leqq 3$, since only free components may have five or more edges. Further $k \geqq 1$, since there are at most four encumbered components, and hence at least two free components.

Subcase 1. $k = 1$ and $f_6 = 0$.

If $\tilde{f}_3 = 0$, then counting dead edges yields

$$8 = \sum e(D) \leqq 1\hat{f}_3 + 2\hat{f}_4 + 2f_5 = 1 + 4 + 2 = 7.$$

Thus this case cannot be optimal when $\tilde{f}_3 = 0$.

If $\tilde{f}_3 = 1$ then $\tilde{f}_4 = 1$, $\hat{f}_4 = 3$, and $f_5 = \hat{f}_5 = 1$. But since two of the four sided encumbered components meet at a vertex of the five-edged component, these two four-faced components can have at most one dead edge. Counting dead edges yields

$$7 = \sum e(D) \leqq 1 + 1 + 2 + 2 = 6.$$

This contradiction shows this subcase cannot be optimal.

Subcase 2. $k > 1$ and $f_6 = k - 1$.

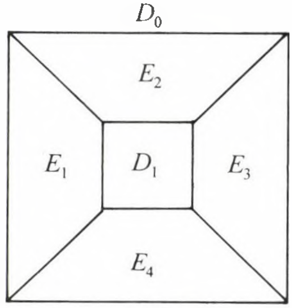

Fig. 14

Since $f_6 > 0$, Lemma 15 tells us that encumbered components have at least four edges, hence $\tilde{f}_3 = 1$, $\tilde{f}_4 = k$, $\hat{f}_4 = 4 - k$, $f_5 = \hat{f}_5 = 1$ and $f_6 = \hat{f}_6 = $ $= k - 1$. Counting dead edges yields

$$3 + 4k = \sum e(D) = 2\hat{f}_4 + 2f_5 + 3f_6 = 2(4 - k) + 2 + 3(k - 1).$$

It follows that $k = 4/3$; so this subcase cannot occur.

Case 8. $(f_3, f_4, f_5) = (1, 3, 3)$.

There are two possibilities:

$$k = 2,\ \tilde{f}_3 \leqq 1,\ \tilde{f}_4 = 3 - \tilde{f}_3,\ \hat{f}_4 = 3 - \tilde{f}_4 = \tilde{f}_3,\ f_5 = \hat{f}_5 = 3 \text{ and } f_6 = \hat{f}_6 = 0;$$

or

$$k = 3,\ \tilde{f}_3 = 1,\ \tilde{f}_4 = 3,\ \hat{f}_4 = 0,\ f_5 = \hat{f}_5 = 3 \text{ and } f_6 = \hat{f}_6 = 1.$$

Counting dead edges for the first possibility,

$$3\tilde{f}_3 + 4\tilde{f}_3 = \sum e(D) \leqq 1\hat{f}_3 + 2\hat{f}_4 + 2f_5 + 3f_6 = 1(1 - \tilde{f}_3) + 2\tilde{f}_3 + 6 + 0,$$

which implies $3 \leqq \tilde{f}_3$; thus this possibility cannot occur. In the second possibility, counting dead edges yields

$$15 = \sum e(D) \leqq 2\hat{f}_4 + 2f_5 + 3f_6 = 0 + 6 + 3 = 9,$$

another contradiction. Case 8 is not possible for an optimal network.

Case 9. $(f_3, f_4, f_5) = (0, 6, 0)$.

In this case $k \geqq 1$, $f_6 = k - 1$, $\tilde{f}_4 = k + 1$, $\hat{f}_4 = 5 - k$, $f_5 = 0$, and $f_6 = k - 1$. Counting dead edges yields

$$4(k + 1) = \sum e(D) \leqq 2\hat{f}_4 + 2f_5 + 3f_6 = 2(5 - k) + 0 + 3(k - 1),$$

it follows that $k = 1$. We are thus led to a network with six four-sided faces, the Schlegel diagram of the cube, Figure 14. There are two possibilities: all edges of D_0 have positive curvature; an edge of D_0 without positive curvature is surrounded by edges of positive curvature.

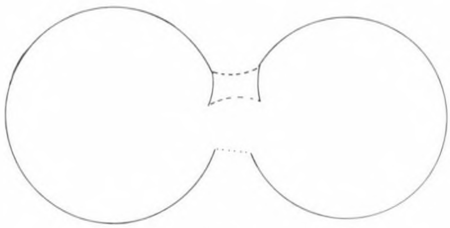

Fig. 15a

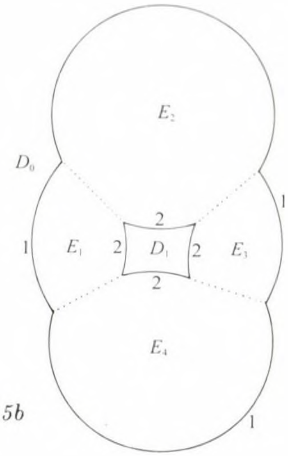

Fig. 15b

The first possibility, Figure 15a, leads to a component with no edges of positive curvature, violating optimality.

In the second possibility, Figure 15b, no edge of D_1 has positive curvature. Case 9 cannot arise in an optimal network.

Case 10. $(f_3, f_4, f_5) = (0, 5, 2)$.

In this case, $k \geqq 2$. It follows that $\tilde{f}_4 = k + 1$, $\hat{f}_4 = 4 - k$, $f_5 = 2$, and $f_6 = k - 2$. Counting dead edges yields

$$4(k+1) = \sum e(D) \leqq 2\hat{f}_4 + 2f_5 + 3f_6 = 2(4-k) + 4 = 3(k-2)$$

it follows that $k < 1$, a contradiction. Case 10 cannot occur in an optimal network.

Case 11. $(f_3, f_4, f_5) = (0, 4, 4)$.

In this case $k = 3$, so $\tilde{f}_4 = 4$ and $\hat{f}_5 = f_5 = 4$. Counting dead edges yields

$$16 = \sum e(D) \leqq 2f_5 = 8,$$

a contradiction. Thus, Case 11 cannot arise in an optimal network.

It remains to consider the curvatures in the two cases of $(f_3, f_4, f_5) =$ $= (2, 3, 0)$ which might arise as an optimal network, neither of which have dead spaces. We first consider the case when $e(D_0) = 3$, Figure 9a. According to Lemma 10, all the edges of D_0 have positive curvature. By symmetry we may suppose that $\kappa_1 \geqq \kappa_2 \geqq \kappa_3 > 0$, where κ_i is the curvature of the edge separating E_i from D_0. If the edge separating E_3 from E_4 does not have positive curvature from the E_4 perspective, then E_4 would have no edges with positive curvature. Thus the only possibilities are shown in Figures 4a (i) and 4a (ii).

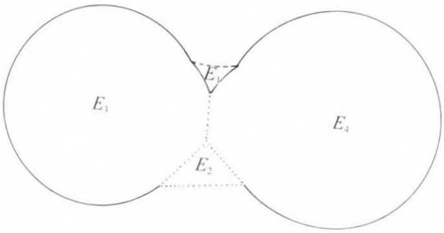

Fig. 16a

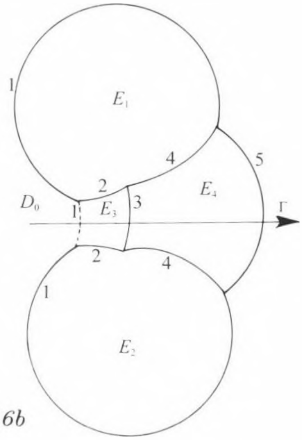

Fig. 16b

If $e(D_0) = 4$, the network has the form of Figure 9b. If a three-edged face does not have positive curvature edge separating it from D_0 and the two four-edged faces do, Figure 16a, then the three edged face has no edges of positive curvature, a contradiction. If both three edged components have positive curvature edges separating them from D_0, but a four-edged face does not, Figure 16b, we are again led to a contradiction as follows: The edges separating E_3 from E_1 and E_2 do not have positive curvature from the perspective of E_3. It follows that the edge separating E_3 from E_4 must have positive curvature from E_3. By looking at the two internal nodes, we conclude that the edges separating E_4 from E_1 and E_2 both have negative curvature from the perspective of E_4. Thus the remaining edge of E_4, separating it from D_0 must have positive curvature. The path, Γ, from D_0 to D_0 crosses only edges of positive curvature, a contradiction to Lemma 7. By Lemma 11 those are the only other possibility if all edges have positive curvature, Figure 4b.

4. Problems and conjectures

The first two are the obvious conjectures to which everyone who has thought about the problem subscribes.

CONJECTURE 1. In an optimal network, there are no dead spaces.

CONJECTURE 2. In an optimal network, each region consists of one component.

The next conjecture is less certain.

CONJECTURE 3. The only configuration possible for an optimal network of four regions is shown in Figure 4b.

The following problem is related, but not equivalent to the last conjecture.

PROBLEM 1. Given a set of parameters can there be more than one optimal network for those parameters, where two solutions are considered the same if they are isometric?

If two different topological types neither of which can degenerate into the other, such as those of Figure 4, appear as optimal networks for different values of the parameters with the same value of n, then for some transitional value of the parameters, there would be optimal networks of two types.

CONJECTURE 4. If the interior of a convex set is divided into n regions of prescribed area which sum to the total area, then the regions are connected in the optimal network.

Probably the hypothesis of convexity is too strong; however if it is relaxed all the way to merely a connected Jordan measurable set, then there can be cases where regions are disconnected, reflecting idiosyncrasies of the boundary. Related work can be found in [4, 8, 9, 13, 14, 21].

PROBLEM 2. What is the appropriate hypothesis to replace convexity in Conjecture 4?

REFERENCES

[1] ALMGREN, JR., F. J., Existence and regularity almost everywhere of solutions to elliptic variational problems with constraints, *Mem. Amer. Math. Soc.* **4** (1976), no. 165. *MR* **54** #8420

[2] BECK, A., BLEICHER, M. N. and CROWE, D. W., *Excursions into mathematics*, Worth, New York, 1969.

[3] BLEICHER, M. N., Isoperimetric division into a finite number of cells in the plane, *Studia Sci. Math. Hungar.* **22** (1987), 123–137. *MR* **89a**:52036

[4] BLEICHER, M. N., Isoperimetric divisions into several cells with natural boundary, *Intuitive geometry* (Siófok, 1985), Colloq. Math. Soc. János Bolyai, **48**, North-Holland, Amsterdam–New York, 1987, 63–84. *MR* **88j**:52023

[5] COX, C., HARRISON, L., HUTCHINGS, M., KIM, S., LIGHT, J., MAUER, A. and TILTON, M., The shortest enclosure of three connected areas in the $\mathbf{R}^2$, preprint dated May 14, 1993 (available from Frank Morgan, Williams College).

[6] EBERHARD, V., *Zur Morphologie der Polyeder*, B. G. Teubner, Leipzig, 1891. *Jb. Fortschritte Math.* **23**, 544–551.

[7] EGGLESTON, H. G., *Convexity*, Cambridge Tracts in Mathematics and Mathematical Physics, No. 47, Cambridge University Press, Cambridge, 1958. *MR* **23** #A2123

[8] FEJES TÓTH, G., An isoperimetric problem for tessellations, *Studia Sci. Math. Hungar.* **10** (1975), 171–173. *MR* **56** #3748

[9] FEJES TÓTH, L., On shortest nets with meshes of equal area, *Acta Math. Acad. Sci. Hungar.* **11** (1960), 363–370. *MR* **23** #A3510

[10] FEJES TÓTH, L., Isoperimetric problems concerning tessellations, *Acta Math. Acad. Sci. Hungar.* **14** (1963), 343–351. *MR* **28** #529

[11] FEJES TÓTH, L., What the bees know and what they do not know, *Bull. Amer. Math. Soc.* **70** (1964), 468–481. *MR* **29** #524

[12] FEJES TÓTH, L., *Regular figures*, Internat. Ser. of Monographs on Pure and Appl. Mathematics, Vol. **48**, Macmillan, New York, 1964. *MR* **29** #2705

[13] FEJES TÓTH, L., *Lagerungen in der Ebene, auf der Kugel und im Raum*, Zweite verbesserte und erweiterte Auflage, Die Grundlehren der mathematischen

Wissenschaften, Band 65, Springer-Verlag, Berlin–New York, 1972. *MR* **50** #5603

[14] Fejes Tóth, L., Isoperimetric problems for tilings, *Mathematika* **32** (1985), 10–15. *MR* **87c**:52023

[15] Foisy, J., Alfaro, M., Brock, J., Hodges, N. and Zimba, J., The standard soap bubble in R^2 uniquely minimizes perimeter, *Pacific J. Math.* **159** (1993), 47–59.

[16] Grünbaum, B., *Convex polytopes*, Pure and Applied Mathematics, Vol. 16, Interscience Publishers [John Wiley & Sons], New York, 1967. *MR* **37** #2085

[17] Hass, J. and Morgan, F., Geodesics and soap bubbles on surfaces (preprint dated September 20, 1991).

[18] Morgan, F., Soap bubbles in $\mathbf{R}^2$ and on surfaces, *MAA Monthly* **101** (1994) (to appear).

[19] Pólya, G., *Induction and analogy in mathematics. Mathematics and plausible reasoning*, Vol. I, Princeton University Press, Princeton, N. J., 1954. *MR* **16**–556

[20] Pólya, G. and Szegő, G., *Isoperimetric inequalities in mathematical physics*, Annals of Math. Studies, No. **27**, Princeton University Press, Princeton, N. J., 1951. *MR* **13**–270

[21] Tomonaga, Y., *Geometry of length and area. 1*, Dept. of Mathematics, Utsunomiya University, Utsunomiya, 1974. *MR* **55** #3977

[22] Weaire, D. and Rivier, N., Soap cells and statics – Random patterns in two dimensions, *Contemp. Phys.* **25** (1984), 59–99.

(Received January 17, 1994)

MATHEMATICS DEPARTMENT
UNIVERSITY OF WISCONSIN
480 LINCOLN DRIVE
MADISON, WI 53706–1388
U.S.A.

bleicher@math.wisc.edu

Studia Scientiarum Mathematicarum Hungarica

Volume 31

1828

Akadémiai Kiadó, Budapest

1996

STUDIA SCIENTIARUM MATHEMATICARUM HUNGARICA

Volume 31

CONTENTS

Typeset by TypoTEX Ltd., Budapest
PRINTED IN HUNGARY
Akadémiai Kiadó & Nyomda, Budapest

RECENTLY ACCEPTED PAPERS

VARMA, A. K., PRASAD, J. and GONSKA, H., On the order of convergence of interpolatory processes

KRISTÓF, J., Strong topology and strong perturbations of normal elements in GW^*-algebras

CSÁKI, E., FÖLDES, A. and RÉVÉSZ, P., Random walk with alternating excursions

IVIĆ, A., The Laplace transform of the square in the circle and divisor problems

YAQUB, A. and HASANALI, M., Commutativity of rings with constraints involving potent and nilpotent elements

ALZER, H., Refinement of a Carleman-type inequality

HERZOG, G., On existence and uniqueness conditions for ordinary differential equations in Fréchet spaces

VERMES, I., Total-separable Kreissysteme und Mosaike in der hyperbolischen Ebene

CSÖRGŐ, M., HORVÁTH, L. and SHAO, Q.-M., Almost sure summability of partial sums

SZYSZKOWICZ, B., Weighted approximations of partial sum processes in $D[0,\infty]$. II

REVENKO, S. and SOLTAN, V., Helly-type theorems on transversality of abstract languages

LEAVITT, W. G., A note on matric-extensibility and the ADS condition

Only original papers will be published and a copy of the Publishing Agreement will be sent to the authors of papers accepted for publication. Manuscripts will be processed only after receiving the signed copy of the agreement.

CONTENTS

HU ISSN 0081-6906